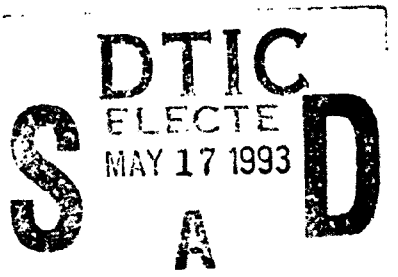
AD-A264 280



RL-TR-92-259
In-House Report
September 1992


# DISPLAY AND ENHANCEMENT OF INFRARED IMAGES

Jerry Silverman and Virgil E.Vickers




*APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.*

Rome Laboratory
Air Force Systems Command
Griffiss Air Force Base, New York

93 5 14 081

This report has been reviewed by the Rome Laboratory Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

RL-TR-92-259 has been reviewed and is approved for publication.

APPROVED: Paul W Pellegrini

PAUL W. PELLEGRINI
Chief, Electromagnetic Device Technology Division
Electromagnetic & Reliability Directorate

FOR THE COMMANDER: Harold Roth

HAROLD ROTH, Director
Solid State Sciences
Electromagnetic & Reliability Directorate

If your address has changed or if you wish to be removed from the Rome Laboratory mailing list, or if the addressee is no longer employed by your organization, please notify RL( ERE ) Hanscom AFB MA 01731-5000. This will assist us in maintaining a current mailing list.

Do not return copies of this report unless contractual obligations or notices on a specific document require that it be returned.

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503

| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | September 1991 | In-house (1988 - 1990) |

**4. TITLE AND SUBTITLE**
Display and Enhancement of Infrared Images

**5. FUNDING NUMBERS**
PE 61102F
PR 2305 TA J1 WU 952680
Grant No. 90-342

**6. AUTHOR(S)**
Jerry Silverman
Virgil E. Vickers

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Rome Laboratory (ESE)
Hanscom AFB, MA 01731-5000

**8. PERFORMING ORGANIZATION REPORT NUMBER**
RL-TR-92-259

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**
Adapted from chapter published in Electro-Optical Displays, edited by Mohammed A. Karim, Marcel Dekker, Inc, New York, 1992

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT** *(Maximum 200 words)*

We review how IR images differ from visible images, emphasize the importance of using a variety of IR images for assessing the utility of various algorithms for displaying them, and develop a gamma function suitable for 8-bit display on Sun workstation monitors. For globally monotonic mapping of wide dynamic range IR imagery to 8 bits for display, we describe and compare direct scaling, histogram equalization, histogram projection, and several hybrid algorithms, including under-sampled projection and plateau equalization. For locally adaptive contrast enhancement, which is not globally monotonic, we describe local histogram projection, modular techniques (sawtoothing), and high frequency sharpening in the spatial and spatial frequency domains. We stress tradeoffs among these algorithms as to degree of enhancement, artifacts, and computational costs. All topics are illustrated by an extensive set of IR images taken with state-of-the-art PtSi Schottky barrier cameras. We also briefly address issues germane to real-time implentation.

| 14. SUBJECT TERMS | | 15. NUMBER OF PAGES |
|---|---|---|
| Infrared imergy | Histogram projection | 82 |
| PtSi cameras | Local contrast enhancement | 16. PRICE CODE |
| Histogram equalization | | |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | SAR |

## TABLE OF CONTENTS




SELECTED S

| Accesion For | |
|---|---|
| NTIS CRA&I | ☑ |
| DTIC TAB | ☐ |
| Unannounced | ☐ |
| Justification | |
| By | |
| Distribution / | |
| Availability Codes | |
| Dist | Avail and/or Special |
| A-1 | |

## ILLUSTRATIONS

## TABLES

# 1. INTRODUCTION

## 1.1 Nature of IR Imagery

High quality infrared (IR) imagery from current staring focal plane arrays has now reached or exceeded TV resolution. For PtSi-based Schottky barrier IR cameras,[1] minimum resolvable temperatures below 0.02 degrees Celsius have been achieved and arrays as large as 512 × 512 are commercially available. Hence, the processing, display and enhancement of high-resolution wide-dynamic-range staring IR imagery, whether for soft-copy display or in real-time hardware embodiments, is becoming an important topic in image processing, but one still in early stages of development.[2]

At the outset, we review the differences between IR images and the more familiar baseline of visible imagery.[3] In visible images, objects reflect the light of a source or sources to a sensor. In ideal (thermal) IR images, the objects emit IR radiation as determined by their absolute temperature; in effect, we view a temperature profile of the scene. However, the monotonic relationship between object brightness and scene temperature can be perturbed by the presence of an IR source such as the sun, for example, which gives a more visual look to daytime IR imagery. Figure 1 contrasts day and night IR images of the same scene – the latter probably closer to an ideal thermal profile. Lesser perturbations in the brightness versus temperature relationship arise from the deviations of real objects from ideal blackbodies as well as from the interaction between the spectral response of the camera and the spectral content of the radiating scene.

Another important distinction lies in the inherently low contrast of IR images compared to visible ones. The typical IR image is dominated by the background radiation at the average scene temperature, as most of the objects will be at ambient temperature and will radiate at roughly the same intensity, leading to a typical variation of a factor of two from lowest to highest signal. Consequently, plots of the number of pixels at various signal levels, the raw image histograms, typically have one high and narrow main peak due to this background radiation (several main peaks if more than one background is present such as sky and ground). Temporal noise as well as small contrast variations within the background will spread out the main peak(s) somewhat but generally leave more concentrated histograms than found for visible images. Frequently, the "targets" or objects of interest are on a small number of pixels and are warmer and hence separated in gray level from the background levels; this leads to a

[1]Shepherd, F. D. (1988). "Silicide Infrared Staring Sensors," Proceedings of SPIE, Orlando, Florida, 930, pp. 2-10.

[2]Silverman, J., Mooney, J. M., and Vickers, V. E. (1990). Display of wide dynamic range infrared images from PtSi Schottky barrier cameras, Opt. Eng., 29: 97.

[3]Silverman, J., Mooney, J. M., and Shepherd, F. D. (1991). Infrared video cameras, Sci. Amer., 266, No. 3, pp. 78-83.


Figure 1. Day and night versions of the same scene.

low level trailing edge in the histogram. The image histograms play a major role in the algorithms of Section 2, and examples and further discussion are found there. A second consequence of the low contrast of IR images is the importance of noise sources, such as spatial noise,[4], which are generally insignificant in the visible.

A subtle point which should be clarified is that this limited contrast does not preclude imagery of such wide dynamic range as often to exceed the 8-bit sensitivity of high quality monitors; and the latter sensitivity is effectively further reduced by the limitations of the inherent gray scale sensitivity of the eye. Images used in this report were taken with PtSi Schottky IR cameras operated in the 3-5 micron band, with noise characteristics which have been carefully analyzed and measured.[4,5] IR image dynamic ranges depend on weather, time of day, detector array technology, camera design, and image content. For PtSi cameras, raw signal levels at the upper end of representative ranges will span about 1000 to 2000 ADUs (analog to digital units) after digitizing single frames to 12 bits. Since a typical noise level is 5 ADUs (see reference 5 for details), a usable dynamic range of up to 200 to 400 levels is often encountered. We are faced therefore with a classical problem in image display: the disparity between the image dynamic range and the smaller dynamic range of the monitor/eye display system.

Ideally, in assessing the relative efficacy of several alternative techniques for display or enhancement, one should work as closely as possible with the type of imagery, the display hardware, the ambient light conditions, etc., specific to the application(s) in question. However, in an imperfect world, the ideal is not always practical. Hence, in evaluating and comparing algorithms for general purpose applications, we have adhered to the following philosophy. We have employed an extensive set of locally taken imagery: indoor, outdoor, day and night scenes as diverse as possible. Whatever general conclusions are advanced in following sections about the relative merits of one algorithm versus another, it is usually not difficult to find an image or image type that belies any particular such conclusion. Therefore, we believe it most useful to emphasize how the various algorithms typically interact with IR imagery. We hope the reader will thereby gain the insight to choose algorithms based on his application and his expected image set.

---

[4]Mooney, J. M., Shepherd, F. D., Ewing, W. S., Murguia, J. E., and Silverman, J. (1989). Responsivity nonuniformity limited performance of infrared staring cameras, Opt. Eng., 28: 1151.

[5]Murguia, J. E., Mooney, J. M., and Ewing, W. S. (1990). Evaluation of a PtSi infrared camera, Opt. Eng., 29: 786.

## 1.2 Display Scales, Algorithms and Applications

With the philosophy just stated in mind, let us turn to the question of the gray scale used for the final 8-bit display, a matter less mundane than one might imagine, as we have found a strong interplay between the display algorithms and the gray scale. The "default" gray scale on our Sun workstation monitors is a linear "colormap" between the display value $i$ and the luminance command value for the red, green and blue components of monitor intensity:

$$\text{red}[i] = \text{green}[i] = \text{blue}[i] = i, \tag{1}$$

where $i$, red, green, and blue all range from 0 to 255. This default scale on our monitors (and we suspect similarly on other monitors) is unbalanced and sub-optimum in that it is too sensitive at the bright end and not sensitive enough at the dark (zero) end. While one could argue that this is desirable for many IR applications where the information of interest tends to be at the hot (bright) end, we would counter that such bias if desired is better introduced into the algorithms rather than into the display scales.

Thus we have assumed that features of interest in the imagery – averaged, so to speak,

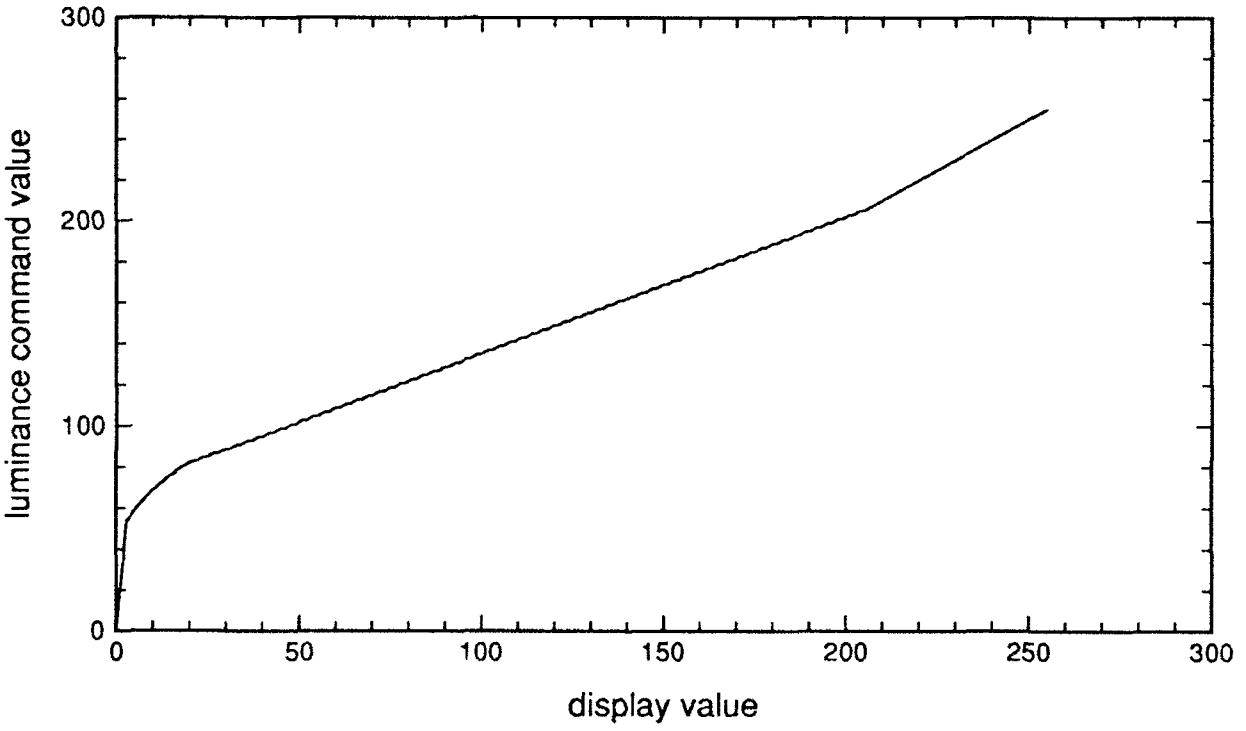



Figure 2. The gamma function.

over many images and applications – are equally likely to occur in any range in the final display. We sought such balance, as well as an optimum number of discernible shades of gray, by going to the form

$$\text{red}[i] = \text{green}[i] = \text{blue}[i] = \gamma(i), \tag{2}$$

where the gamma function for the monitor[6] represents the desired mapping from display value to screen luminance command value.

While formal algorithmic procedures for determining $\gamma(i)$ are available,[7] we have found the following simple procedure adequate. Using software-generated standard bar patterns with a range of spatial frequencies, we set display contrasts at $\Delta i = 2$ or 3. The bar pattern backgrounds were set from dark to light in gradual increments. Working in a darkened room (which means strictly speaking that the scale should always be used in a darkened room – needless to say, it wasn't), one author mapped the above display contrasts into luminance command contrasts that were comfortably perceptible for the larger bar patterns and just perceptible for the smallest pattern. The gamma function thus derived is shown in Figure 2 (the optimum mapping varies slightly from monitor to monitor). Figure 3 shows the difference this gamma function makes, compared with the default gray scale, using both a real daytime image and a simulated image of uniform blocks going from 0 to 255 in unit steps. Note the difference in balance and sensitivity between the scales.

For the gamma-corrected scale, when $i$ is between 35 and 175, a change of 2 is just perceptible, while in the regions above and below these limits, a change of 3 is required. We estimate that roughly 110 shades of gray are discernible out of the 256 nominal levels. Although the use of pseudocolor is beyond the scope of this report, we note in passing that a color scale with about 200 discernible levels has been designed for use with our IR images.[2]

[6]Briggs, S. J. (1987). "Soft Copy Display of Electro-optical Imagery," Proceedings of SPIE, 762, pp. 153-170.

[7]Briggs, S. J. (1981). Photometric technique for deriving a 'best gamma' for displays, Opt. Eng., 4: 651.

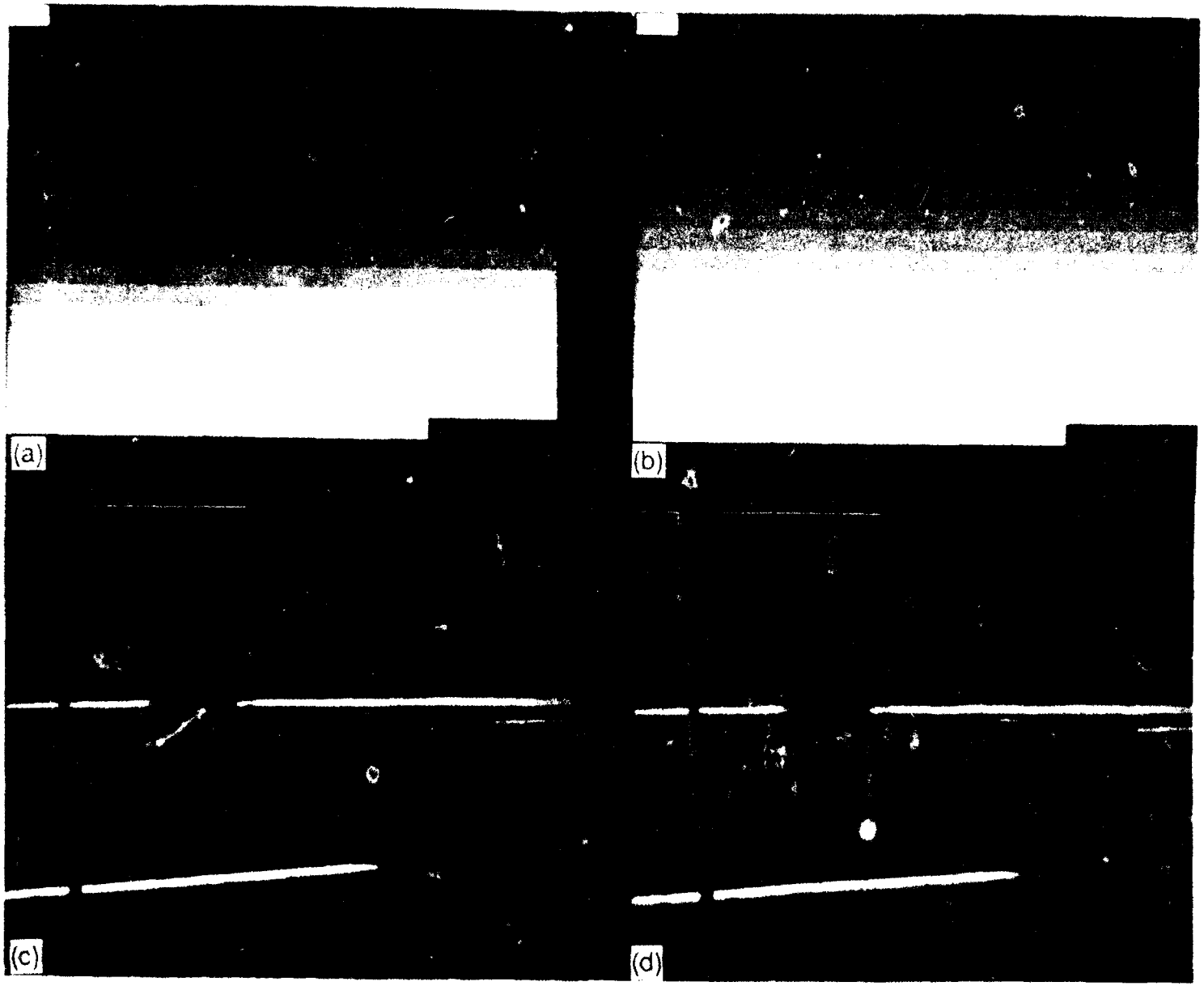



Figure 3. Examples showing effect of gamma function correction: (a), (c) displays with default gray scale; (b), (d) with gamma function correction.

## 2. GLOBAL MONOTONIC DISPLAY ALGORITHMS

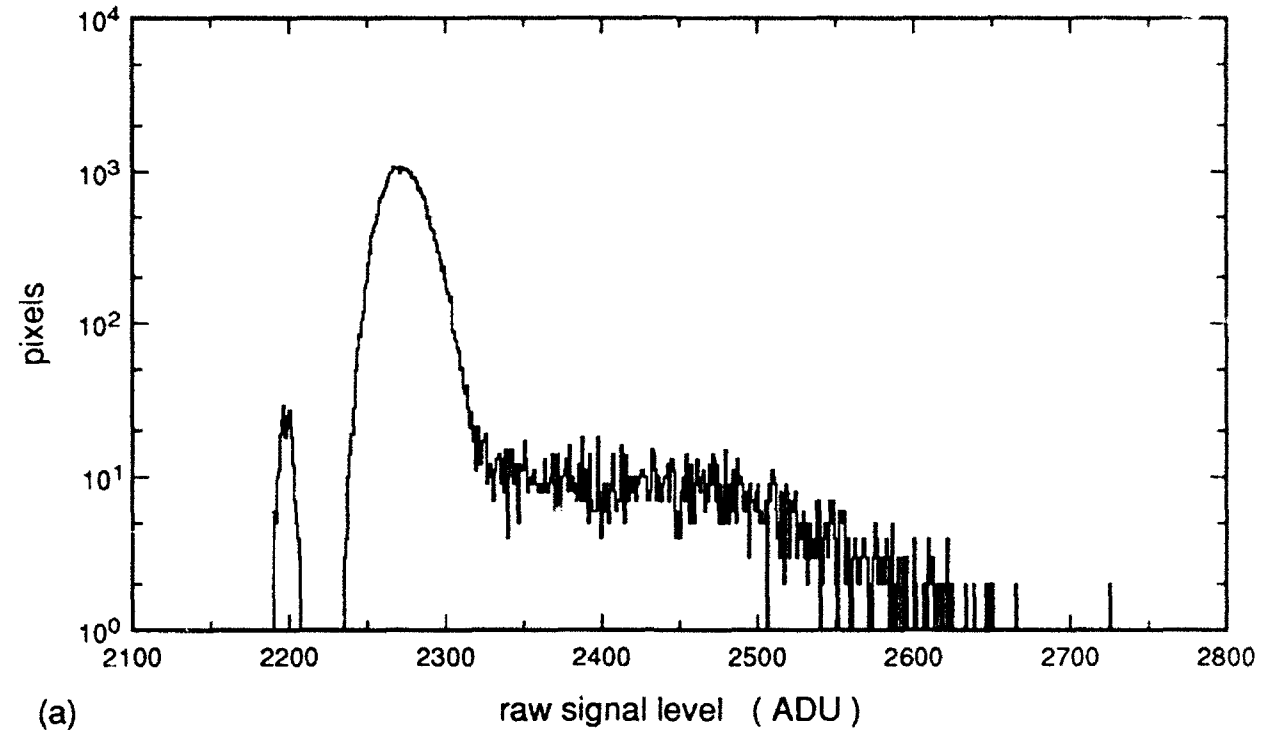
### 2.1 Introduction

To display our images, we need to map from the raw recorded signal (digitized to 12 bits for the images used to illustrate this report) to 8-bit values. Algorithms for this purpose can be divided into a global monotonic group treated in this section, and all others treated in Section 3; Table 1 lists the major algorithms considered in this report. More specifically, we now consider global mappings (no influence of local context) in which the radiometric trend from low to high in the recorded image is retained in the displayed image (monotonic). We found the distinction of algorithm type between Sections 2 and 3 meaningful for IR images; it is typically ignored for visible images, for which raw signal levels depend strongly on natural and artificial light sources in the vicinity[8] (see Fig. 1 of reference 8), and for which the intimate familiarity of the human brain with such imagery allows for flexible interpretation, so that we are not disturbed by deviation from monotonicity.

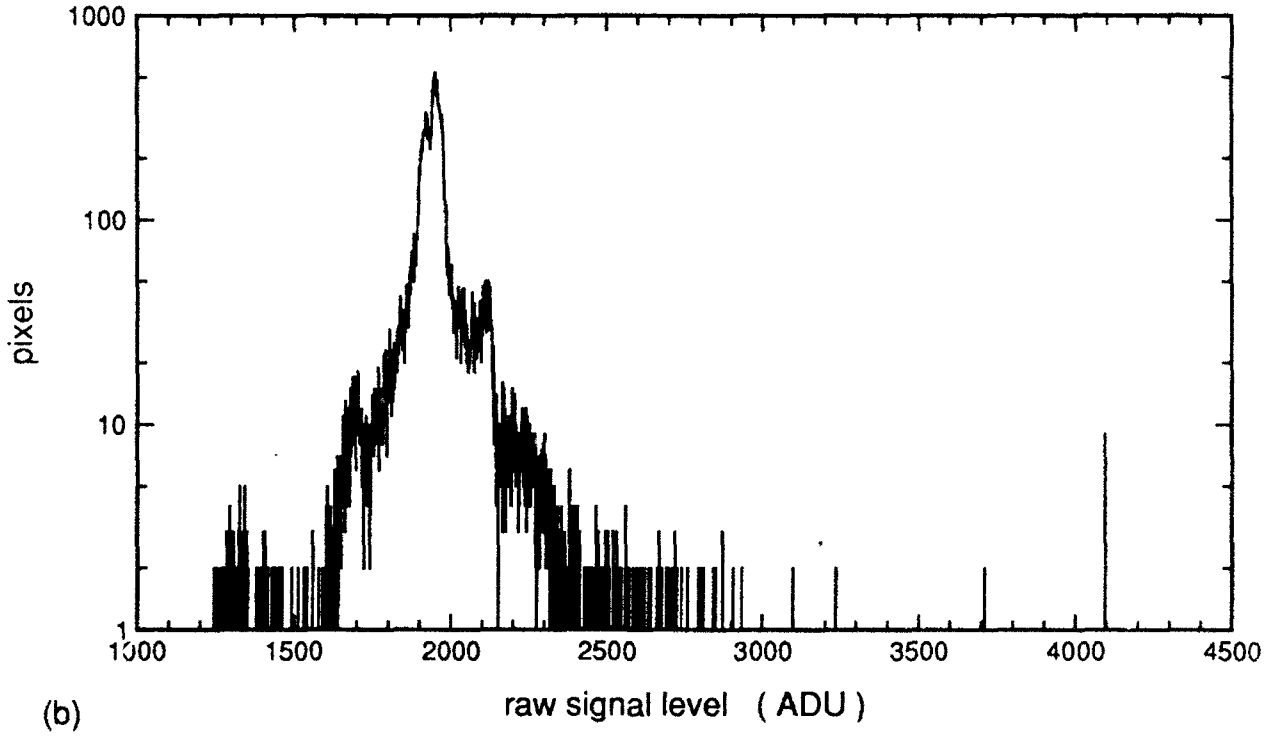
**Table 1. Acronyms of Algorithms Considered**

| Algorithm | Acronym | Section Introduced |
|---|---|---|
| Direct Scaling | DS | 2.2 |
| Histogram Equalization | HE | 2.3 |
| Histogram Projection | HP | |
| Under-sampled Projection | UP | |
| Threshold Projection | TP | |
| Plateau Equalization | PE | |
| Local Range Modification | LRM | 3.2 |
| Overlapping Projection | OP | |
| Sliding Projection | SP | |
| Raw Modulo | RM | 3.3 |
| Modulo Projection | MP | |
| Weak Sinc Sharpening | WS | 3.4 |
| Strong Gaussian Sharpening | SG | |
| Medium Gaussian Sharpening | MG | |
| Weak Gaussian Sharpening | WG | |

[8]Schreiber, W. F. (1978). Image processing for quality improvement, Proc. IEEE, 66: 1640.


pixels
$10^4$
$10^3$
$10^2$
$10^1$
$10^0$
2100
2200
2300
2400
2500
2600
2700
2800
raw signal level ( ADU )

(a)


pixels
1000
100
10
1
1000
1500
2000
2500
3000
3500
4000
4500
raw signal level ( ADU )

(b)

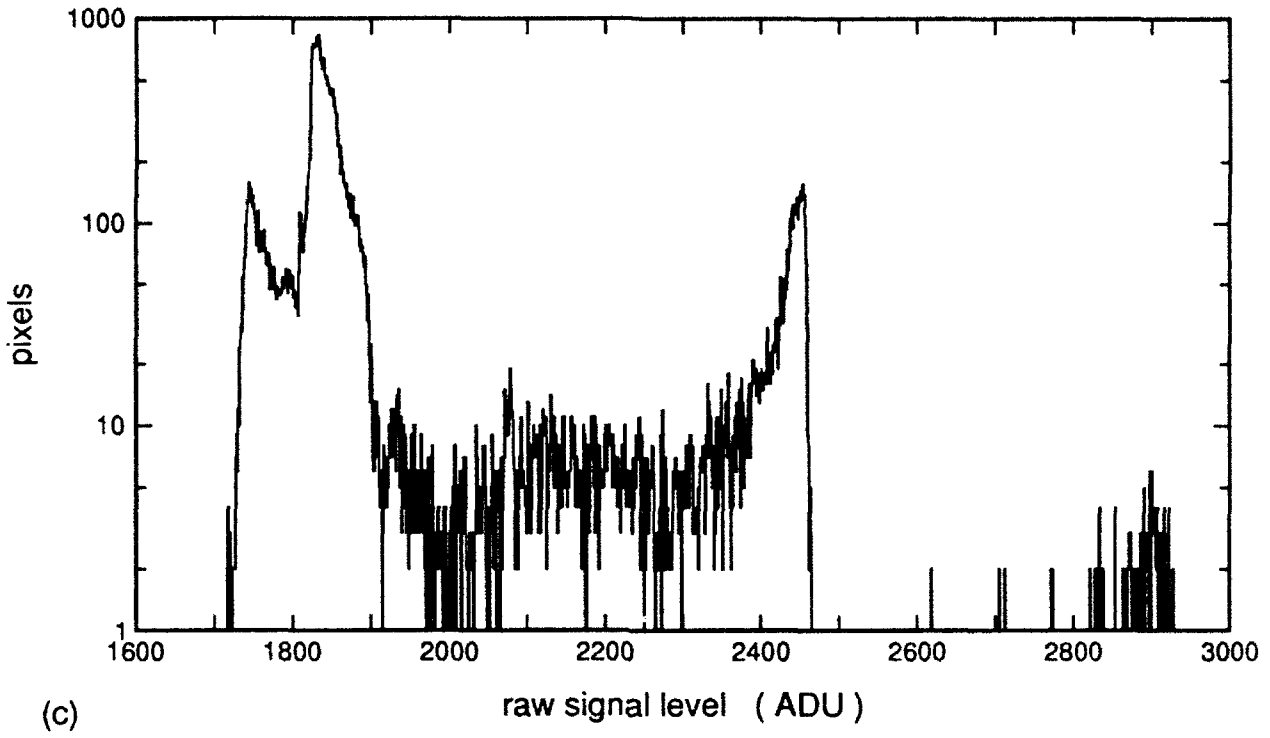



Figure 4. Raw signal histograms for the three standard images: (a) geese; (b) airport; (c) cups.

These global monotonic algorithms are subdivided into direct scaling and related algorithms and histogram-based algorithms. The fundamental distinction here is between the linear or piecewise-linear mapping of the scaling algorithms versus the nonlinear mapping of the histogram-based algorithms. In the former one is reserving dynamic range in the display for empty regions, if present, within the span of the raw signal histogram; while in the latter unoccupied levels are "squeezed" out of the display.

We next introduce a basic image set chosen to illustrate the operation of the various display algorithms. A common practice, particularly in the literature on restoration and for visual imagery, is to employ some familiar standard images previously used by other workers. Aside from the practical matter of the absence of such standard IR images, we reiterate that specific images can be quite misleading or at least unrepresentative of general trends. For example, we firmly contend that histogram equalization, a standard technique discussed in Section 2.3, is an ineffective algorithm for IR images. Yet it is not hard to find images for which the technique is quite satisfactory.

How then does one handle this difficulty and still retain coherence and some degree of brevity? We ask the reader to accept on faith that the small set of three images which will be

used as a common thread throughout the rest of the report to illustrate trends and conclusions are fairly representative of the range of possibilities for the (literally) hundreds of IR images used to test the many algorithms. Where needed to make a special point, or hopefully to prevent boredom, other images will be interjected at certain junctures.

The three images are a sunny day image of Canada geese on a grassy background, a fairly complex night image of an airport scene, and a staged indoor scene of wide dynamic range with hot and cold cups (low contrast details on each cup), between which is a set of bar patterns. The digitized raw signal histograms of the image set are presented in Figure 4. The geese image exemplifies a very common type of IR image with a histogram much like the prototype described in Section 1.1: a very concentrated main peak from the grassy background and a trailing edge on the high side of the peak arising from warmer objects that occupy fewer pixels. The airport scene is a representative night image for mild clear weather (a noisier cold night landscape scene will be used as well). The complex histogram of the two-cup image has portions related to the cold cup (ice water), bar patterns, background, and warm cup (hot water). This staged image is rich in specific local details, and is interesting for another reason as well. The first two images are typical of most of our surveyed images in that they are "grabbed" single frames which were one-point-corrected by camera electronics. The noise of such images includes the temporal noise of the single frame and the residual spatial noise associated with an imperfect correction. The two-cup image is produced from three direct (uncorrected) images; each image is an average over 256 successive measured frames and hence has negligible temporal noise. The three direct images are high and low temperature uniform scenes and the direct cup scene itself, and the final image is the two-point-corrected result,[5] which has lower residual spatial noise than one-point-corrected images.[4] Hence, this image has low contrast details and very little noise. (It will be compared later with a single-frame, one-point-corrected similar image.) The portion of its histogram above 2600 is an artifact of the interaction between the two-point correction and the bad pixels on the top few rows and contains no information, but it can affect the operation of display algorithms.

## 2.2 Direct Scaling and Related Linear Mappings

Direct scaling (DS), if carried out interactively in software, is similar to manual offset/gain adjustment (contrast and brightness) of "live" camera imagery as guided by the eye. In the software interactive mode, the observer views the histogram and chooses a black and white level whose span is then linearly mapped into the full display dynamic range. By and large, the obvious level choices usually provide a display very similar to that given by the histogram projection algorithm discussed in the next subsection, despite the linear/nonlinear distinction between the two mappings. In some cases, the optimum choices are not so obvious, as in the direct-scaled displays of the two-cup image in Figure 5, which take white as 2500 or 2900 respectively. The second choice wastes display dynamic range with no increase

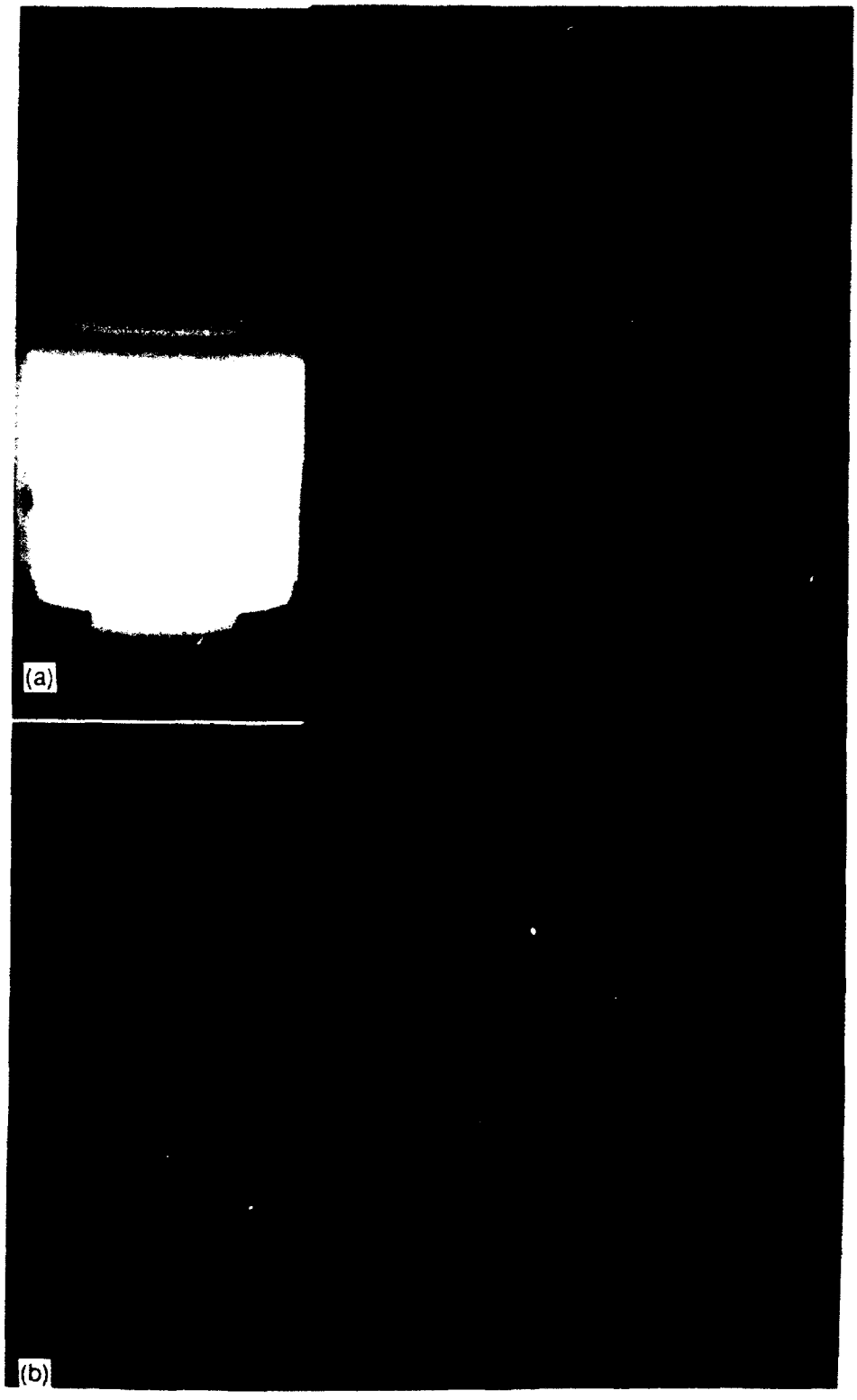



Figure 5. Direct-scaled display with two choices for white level: (a) 2500; (b) 2900. Cf. Fig. 15.4c.


Figure 6. Examples of DS displays with various choices for black/white levels: (a) high/low; (b) 0.05%; (c) 0.1%; (d) 1.0%. See text.

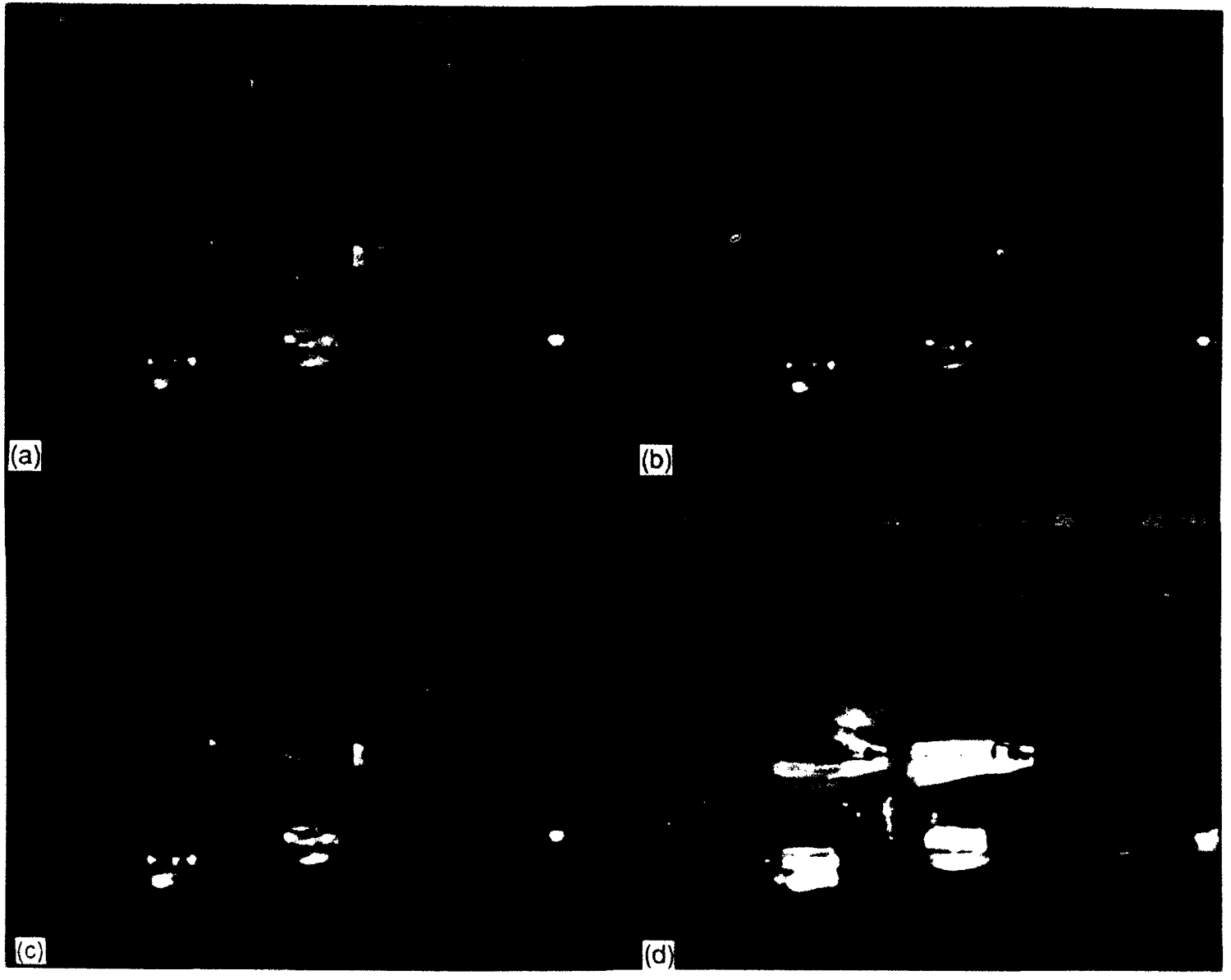



Figure 7. Examples of DS displays with various choices for black/white levels: (a) high/low; (b) 0.05%; (c) 0.1%; (d) 1.0%.

in scene content.

However, a more fundamental problem with DS, particularly for hardware implementation as an automatic offset/gain control to replace manual adjustment, is the difficulty of finding a robust automatic counterpart of the interactive process, because of scene and application dependencies. One simple candidate technique is to determine the black and white levels with symmetrical criteria by "integrating" to some fraction of the total area of the histogram plot from the bottom (black) or top (white) end. Figures 6 and 7 show the results of such an algorithm for the geese and airport images with choices of 0.05%, 0.1% and 1.0% respectively of the area (values are percents of the pixels below the black level and above the white level). Also included is the computationally simple but naive choice in which the lowest and highest occupied raw signal levels are taken as black and white respectively: the presence of a few unreliable pixels typically makes such a high/low scaling a poor choice. The optimum of the integration procedure is often at the 0.1% level; going beyond the optimum, as in the fourth picture of the sequences of Figures 6 and 7, tends to increase overall contrast but leads to a too high black level or too low white one, giving poor gray scale resolution at the low or high end (the latter here).

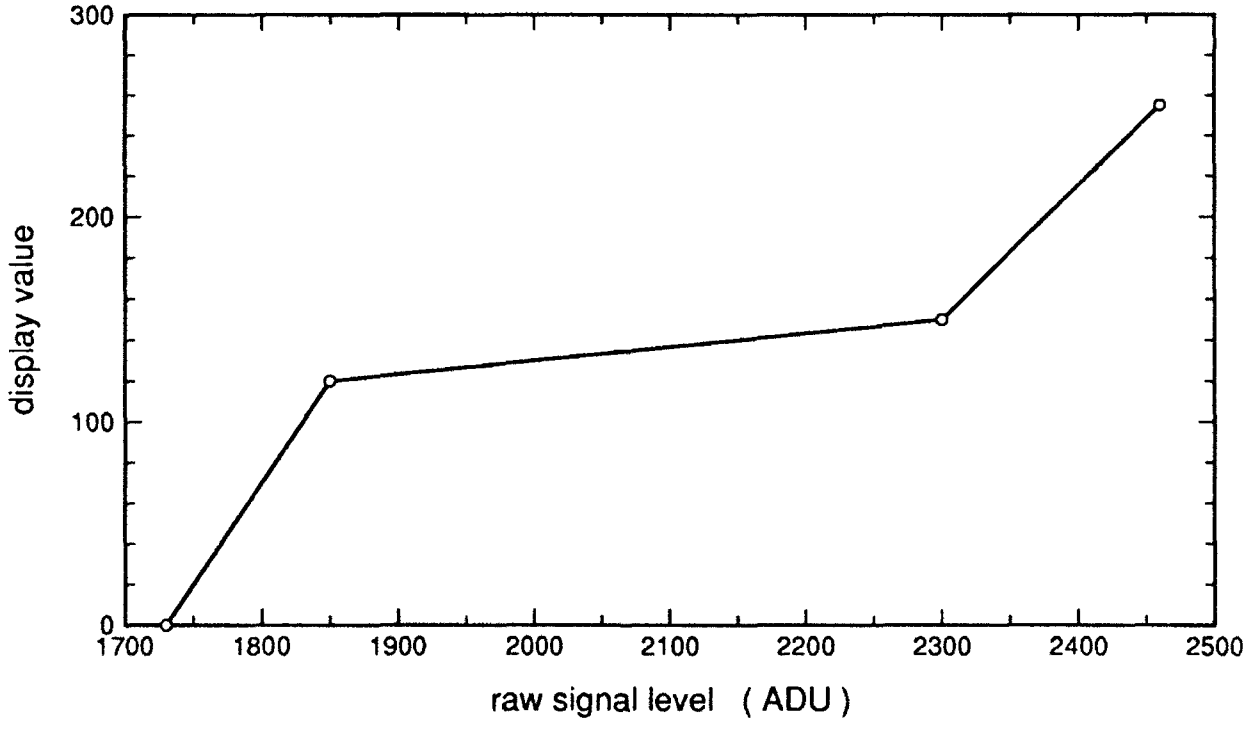



Figure 8. Custom piecewise-linear mapping function for cups image.

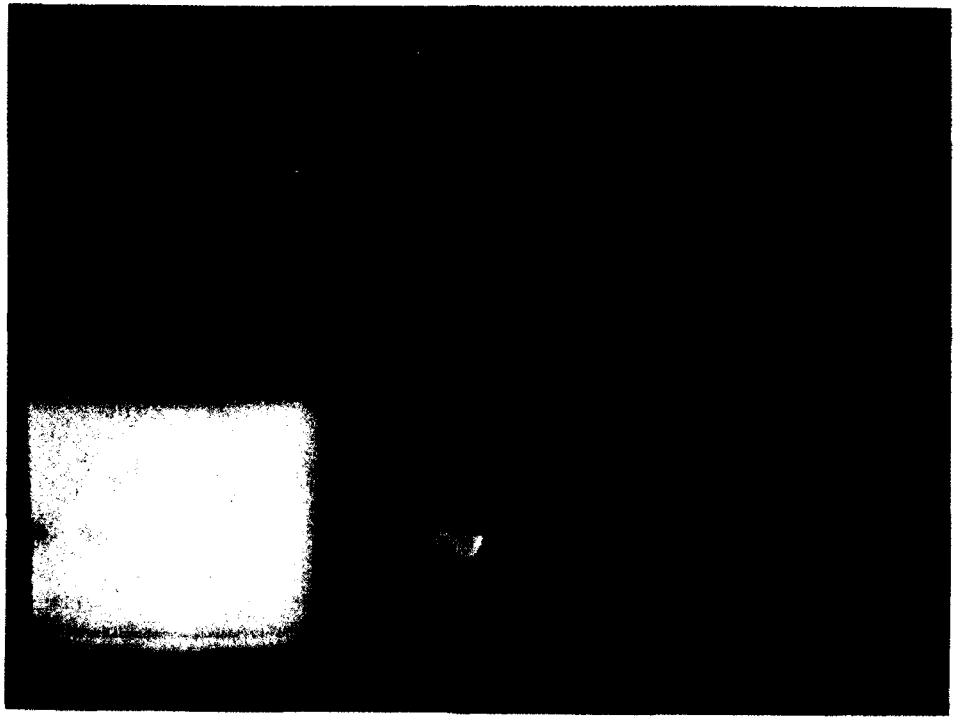


Figure 9. Display based on piecewise-linear mapping shown in Fig. 8.

To increase the utility and robustness of a DS algorithm, particularly for hardware implementation, one would need to incorporate a more sophisticated analysis of the histogram and to allow for unsymmetrical criteria for black and white choices. In our opinion, the results would still fall short in robustness for many image types, when compared with the results of the algorithms described in the next section.

A further extension of the DS concept would be a piecewise-linear transformation, for example of the interactive type referred to as "function processing"[9] but restricted to monotonic mappings for the purposes of this section. Carefully probing the raw signal data of the two-cup image and after some trial and error attempts, we came up with the 3-piece linear mapping shown in Figure 8, affording the display in Figure 9. This is superior to any of the other global monotonic algorithms on this image, although the hybrid algorithms described in the next subsection come close. Since this technique is customized to each image and requires substantial operator intervention, we submit that it is more of a "process" than an "algorithm" and not practical for general hardware or software use.

### 2.3 Histogram-based Nonlinear Mappings

By histogram-based algorithms, we refer generally to nonlinear mappings governed by transformations from the raw signal histogram to some desired final display histogram. The prototype for such methods is the well-known technique of histogram equalization (HE) described in many texts on image processing.[10] As the name indicates, the desired final form is a uniform histogram distribution. While several variations of HE have been proposed,[11] including hyperbolic and exponential distributions as desired goals, we believe such refinements are essentially related to shifts in the gamma function of the gray-scale display (Section 1.2) which can be treated at a "pre-algorithmic" stage. Hence, our treatment in this section will focus on HE and a newer polar opposite to it called "histogram projection" (HP), as well as hybrids of the two.

HE often gives excellent results on visible imagery and is claimed to be optimum from the standpoint of information theory.[12] Generally it has been used to redisplay 8-bit data on an 8-bit scale. For the present purpose, mapping 12-bit IR data to 8 bits, the algorithm has major problems.

---

[9]Woods, R. E., and Gonzalez, R. C. (1981). Real-time digital image enhancement, Proc. IEEE, 69: 643.

[10]Pratt, W. K. (1978). Digital Image Processing, John Wiley & Sons, New York, pp. 307-344.

[11]Hummel, R. (1977). Image enhancement by histogram transformation, Comp. Graph. & Imag. Proc., 6: 184.

[12]Tom, V. T., and Wolfe, G. J. (1982). "Adaptive Histogram Equalization and its Applications," Proceedings of SPIE, 359, pp. 204-209.

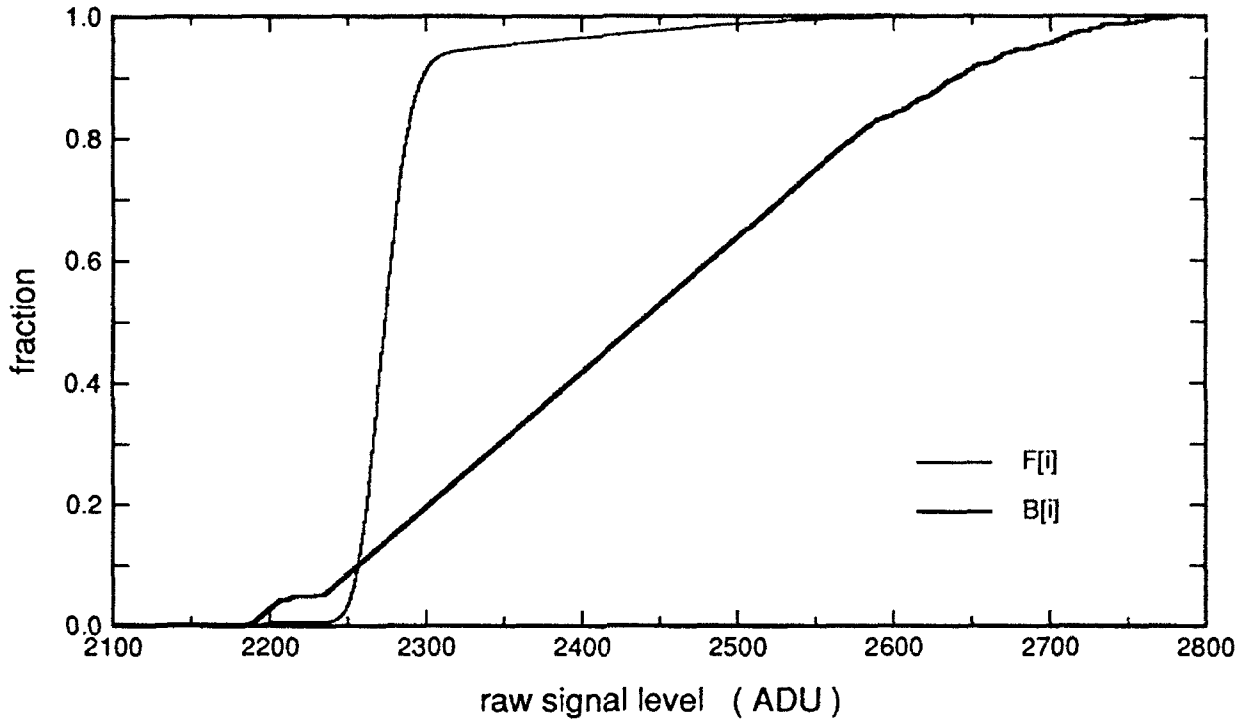



Figure 10. Cumulative distribution functions for geese image.

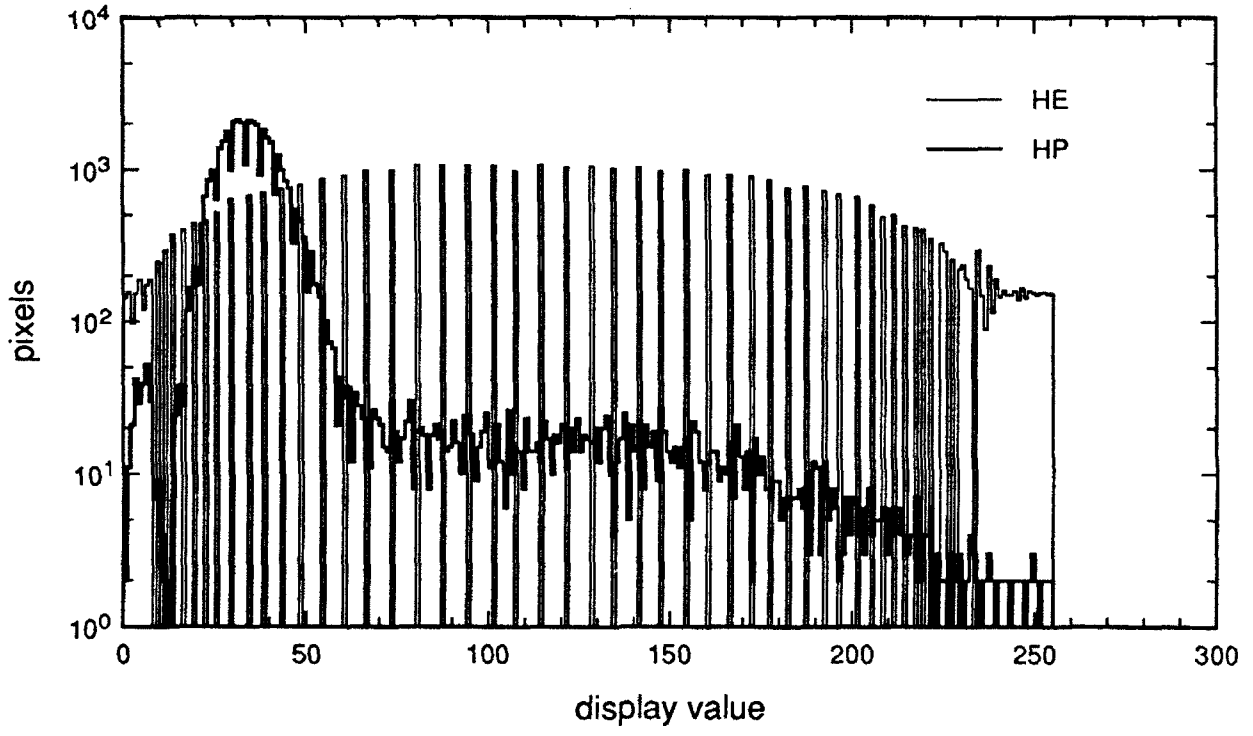



Figure 11. Display histograms for HE and HP for geese image.

Figure 12. Comparison of HE and HP displays: (a) HE; (b) HP.

To implement HE for the discrete case, one converts the histogram of the starting (raw) data to a cumulative distribution function, $F[i]$, which rises in discrete jumps from 0 to 1 and specifies what **fraction of pixels** are at or below the raw signal level $i$. Figure 10 shows $F[i]$ for the geese image. Note the large jump in $F[i]$ within the histogram peak (refer back to Fig. 4a). For display on an 8-bit scale, pixels at raw level $i$ are mapped into

$$\text{display value} = 255\, F[i]. \tag{3}$$

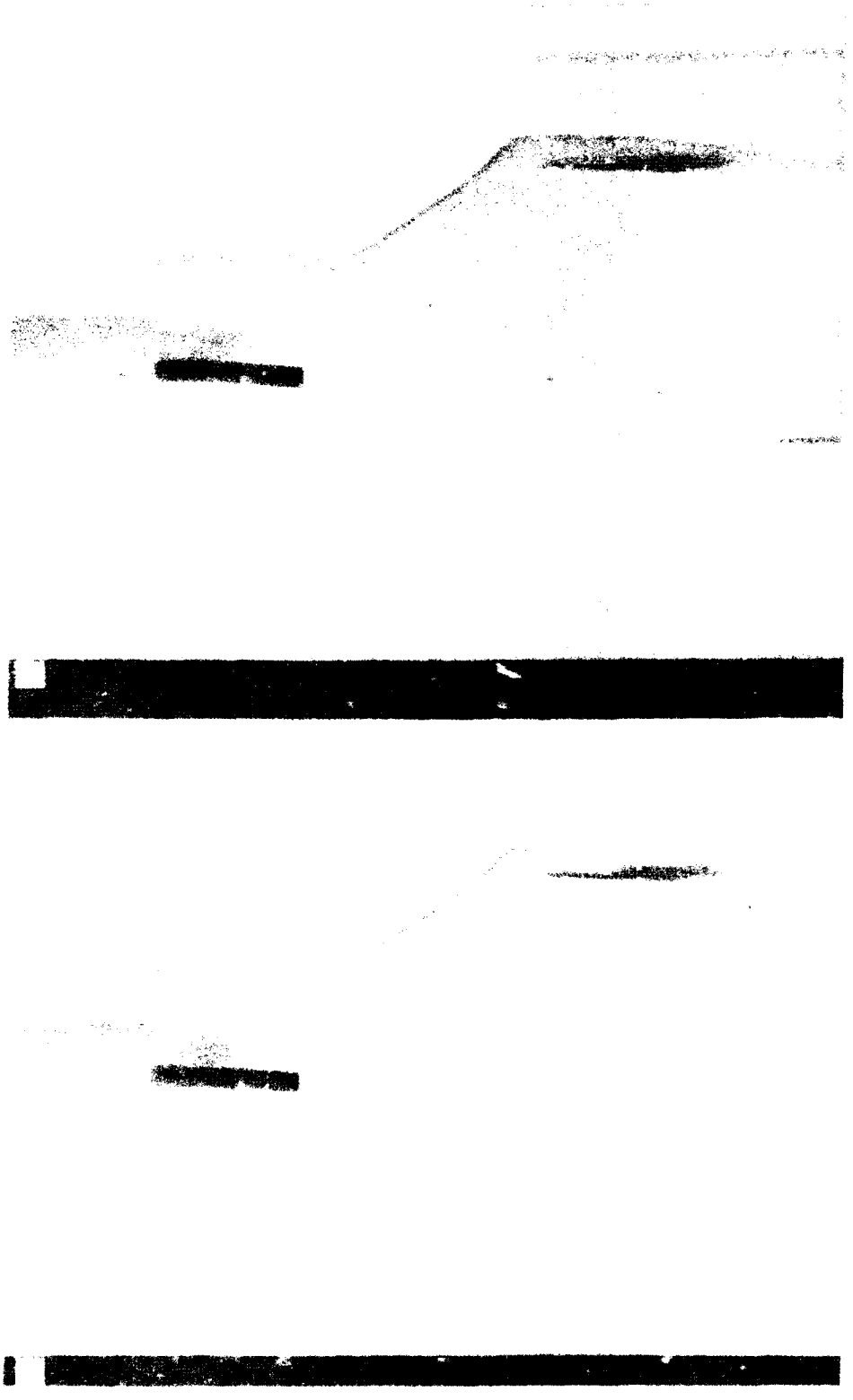
The resulting display histogram and displayed image are shown in Figures 11 and 12 respectively. The transformation of Eq. 3 produces an approximately uniform display histogram by fusing sparsely-occupied adjacent raw signal levels while reserving more display dynamic range for signal levels with high pixel counts (places where $F[i]$ has large jumps). Empty levels can even be created on the display scale (Fig. 11) between densely-occupied adjacent raw signal levels. Referring back to the starting histogram (Fig. 4a), we see that the small leading peak at 2200 and the long trailing stream beyond 2300 retain their identity in the display histogram over a very narrow display range at the dark and light ends.

We reiterate that the geese scene is representative of a common type of IR image in which most of the pixels are on a background (ground here) with relatively little detail or variety, while a minority of pixels are on smaller objects of interest whose signal levels are separated from this background. As a measure of the degree of histogram concentration (Fig. 4a), we note that 25% of the total of 451 occupied raw signal levels (the presence of at least one pixel at a given signal level defines occupancy) account for 95% of the pixels. For images of this common type, HE assigns most of the display levels to the small variations in the background and its associated noise. Raw signal levels of the small objects, if outside this background peak, are compressed into few display levels. (Hence the geese are displayed as white blobs starkly separated from the background but without the internal thermal detail contained in the raw data.)

This incompatibility between IR images and the HE technique was pointed out by Dion and Cantella[13] as follows: "since it operates on the basis of probability of occurrence, a small object within a given field-of-view can ordinarily be de-emphasized if its detail lies at an infrequently occurring level". Their solution was to precede HE with a high-pass filter. However, one can then no longer guarantee a global, monotonic mapping, and further the tendency of HE to amplify the background noise is increased. Another solution, first reported in 1988,[14] is an algorithm, "histogram projection" (HP), whose guiding principle is

[13]Dion, D. F., and Cantella, M. J. (1984). Real-time dynamic range compression of electronic images, RCA Eng., 29: 42.

[14]Silverman, J., and Mooney, J. M. (1988). "Processing of IR Images from PtSi Schottky Barrier Detector Arrays," Proceedings of SPIE, San Diego, California, 974, pp. 300-309.



diametrically opposite to HE: display dynamic range is assigned equally to each signal level present, regardless of how many pixels occupy that level.

To perform HP, one need only compute an occupancy (binary) histogram and then order the occupied raw signal levels from 1 to $N$ from lowest to highest with $N$ the total number of levels occupied. The cumulative distribution corresponding to $F[i]$, now called $B[i]$, represents the **fraction of occupied levels** at or below the level $i$. $B[i]$ rises from 0 to 1 in discrete uniform steps of $1/N$ and is also illustrated in Figure 10.

An expression similar to Eq. 3 may be used to describe the 8-bit display of HP, namely

$$\text{display value} = \lfloor 256\,(B[i] - 1/N) \rfloor, \tag{4}$$

where $\lfloor\ \rfloor$ represents truncation to the next lower integer. The uniform step rises in $B[i]$ in effect assign equal display space to each occupied level. A more accurate description of HP may be based on thinking of the $N$ occupied levels as linearly mapped or "projected" into the 8-bit display scale. If $N$ is greater than 256 levels, neighboring occupied levels in the raw signal are fused to the same display value on the 8-bit scale by the compression factor, $N/256$. In a typical software implementation using integer arithmetic, one could write for each pixel

$$\text{display value} = \lfloor 256\,(n - 1)/N \rfloor, \tag{5}$$

where $n$ is the order number (from 1 to $N$) of the pixel's occupancy level.

The display of the geese image according to HP is also seen in Figure 12 and the corresponding display histogram in Figure 11. The close correspondence between this display histogram and the original (Fig. 4a) is typical of the algorithm. Indeed, thinking in terms of a transform driven by a desired final histogram, one seeks with HP to project the original histogram (excising empty levels) into the available display space – corresponding features such as peaks become higher, of course, if $N$ is greater than 256. The natural, although somewhat dark-level, view of the background and the excellent resolution in gray-scale of the smaller, warmer objects are characteristic of this algorithm.

The HP display is typically indistinguishable from the best DS result, despite the nonlinear/linear difference between the two mappings, as in the comparison of Figure 13 for the airport image (the keen-eyed viewer might spot some differences such as in the airplane windows). With the increasing sensitivity of IR imagers, the influence of unoccupied levels within the linear span of the image signal content may well become more important and the payoff from the nonlinear feature more apparent. In any case, the HP algorithm is easier to implement in real-time and more robust than the DS algorithm.

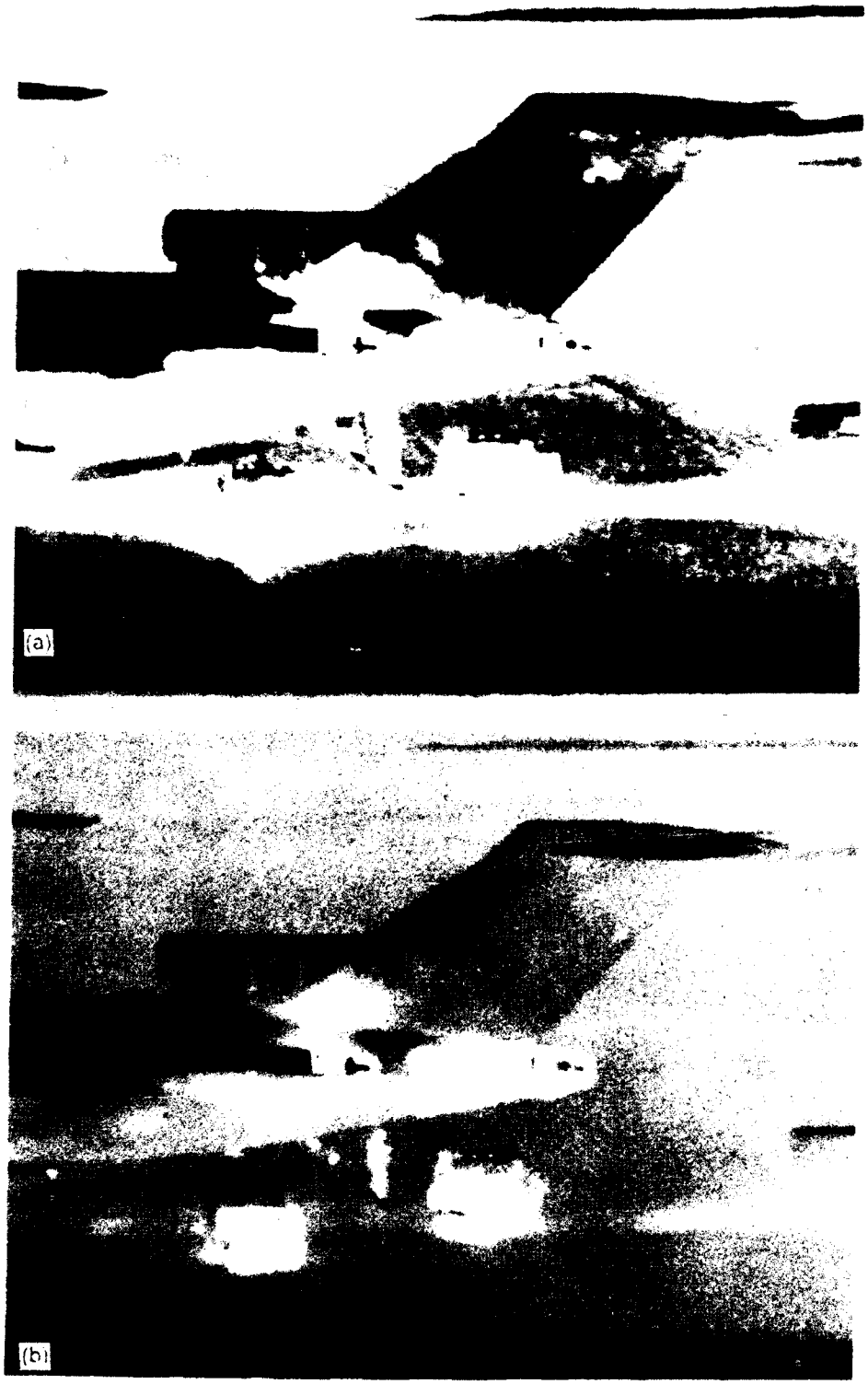



Figure 14. Comparison of HE and HP displays: (a) HE; (b) HP.

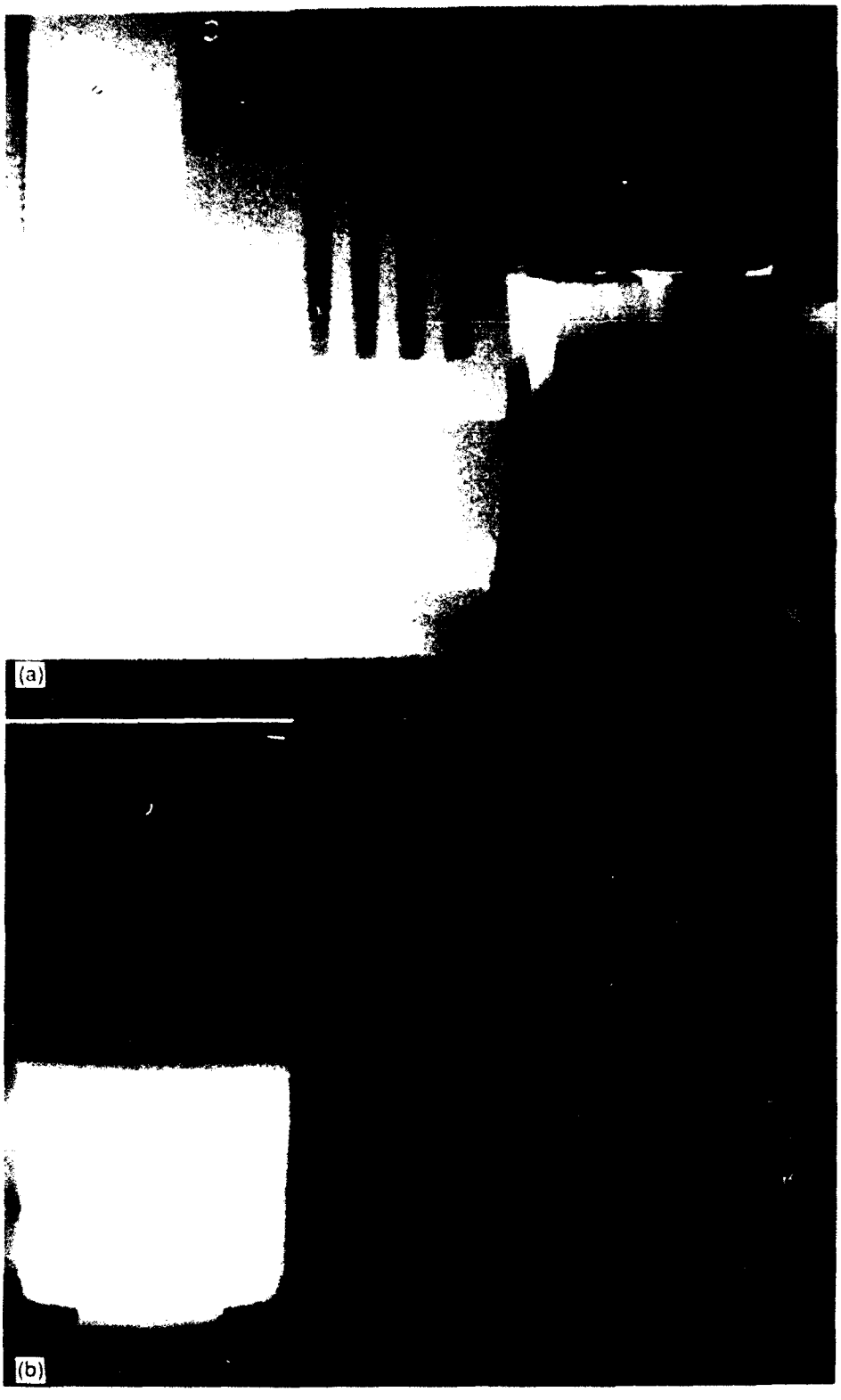



Figure 15. Comparison of HE and HP displays: (a) HE; (b) HP.

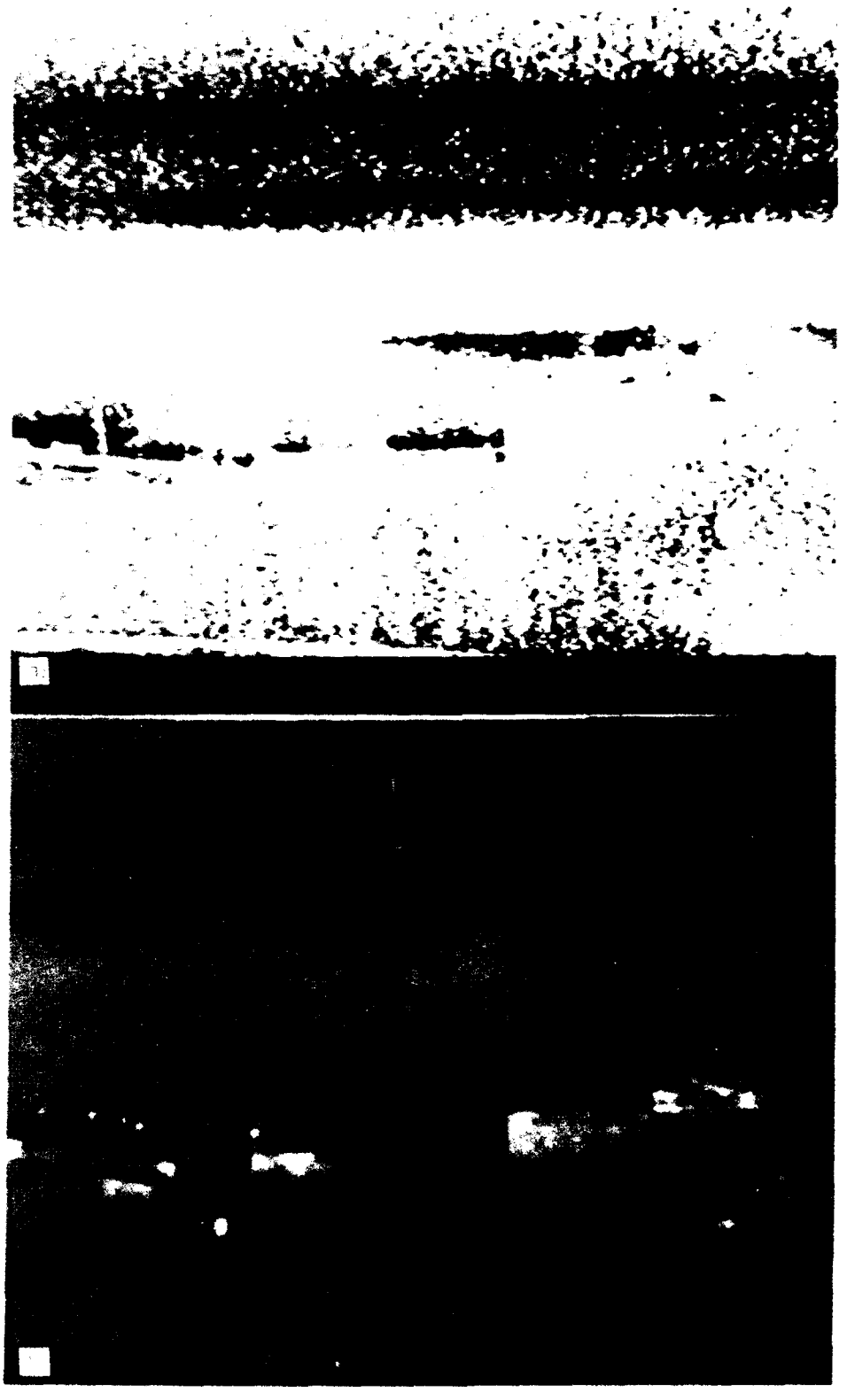


Figure 16. Comparison of HE and HP displays: (a) HE; (b) HP.

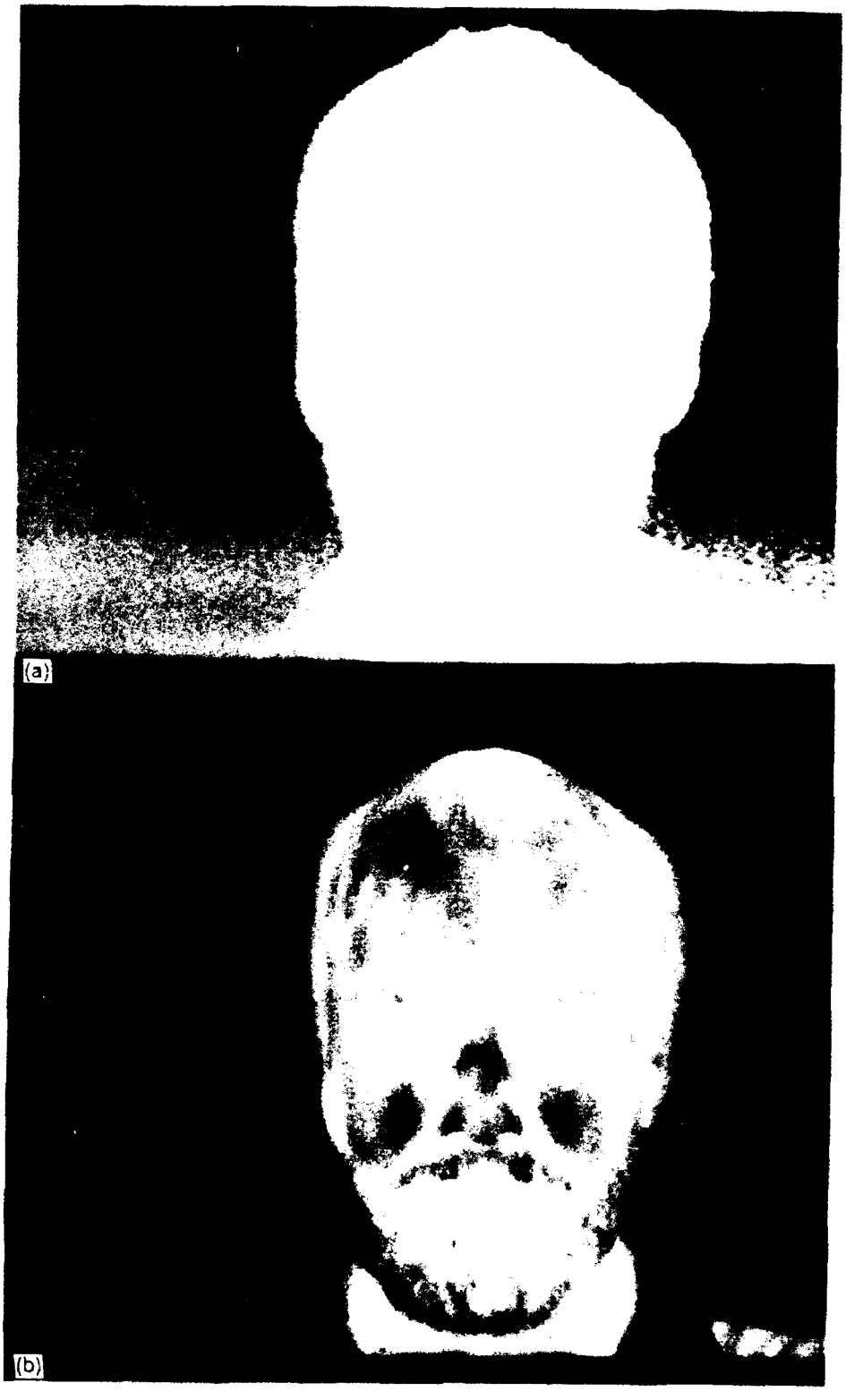



Figure 17. Comparison of HE and HP displays: (a) HE; (b) HP.

Some further comparisons of the displays generated by HE and HP are given in Figures 14-17. For the airport image, the displays are more complementary in their strong and weak points with the usual superior gray-scale resolution of small-scale objects in the HP result (the person and the vehicle grilles) but with a broader delineation and hence clearer spatial sense between foreground and background in the HE display. The face tends to favor the HP result, but if we had less neutral backdrop and more face in the field-of-view, the two displays would be more comparable. The cold winter-night landscape (Fig. 16) is a striking example of the interaction of temporal noise at the background levels with the two algorithms: the tendency of HE to amplify such noise is horrendous here, to the extent that image recognition is almost destroyed. Finally, the two-cup image displays represent one of a small number of cases where HE gives the better overall display. In these instances, some hybrid form of the two algorithms is usually superior to either; this brings us to the next topic in this subsection.

The allocation of display dynamic range according to histogram height, the prominent effect of HE, can be beneficial when used in a weaker mode than occurs in HE. Images that especially benefit from such weighting, such as the two-cup image, have rather complex multi-peaked histograms and little information of interest in sparsely-occupied raw histogram levels. For example, as described above, the occupied levels above 2600 in the histogram of the cup image (Fig. 4c) are artifacts. These "eat up" dynamic range in the HP display.

Our first attempt to fuse the two algorithms was literally a "hybrid" process in which a weighted combination of the cumulative distribution functions of each is used:

$$\text{display value} = W \lfloor 256\,(B[i] - 1/N) \rfloor + (1 - W)\,255\,F(i). \qquad (6)$$

Values of $W$ from 0.9 to 0.7 often give the best result, i.e., "mixing-in" between 10 and 30% of the HE weighting effect. Figure 18 shows the displays resulting from a 25/75% mix of HE/HP for the airport and two-cup images. Since such a hybrid procedure reintroduces and even accentuates the computational complexity of the HE algorithm, we sought alternatives which hybridize the results without such increased complexity. Three are discussed here: two are basically variations on HP called "under-sampled projection" (UP) and "threshold projection" (TP), and one is really a variation on HE, "plateau equalization" (PE).

Before examining these three algorithms, one should emphasize that each of them depends on a single parameter which introduces the assignment of dynamic range on the basis of histogram height in a gradual and controlled manner. As implemented in software on actual imagery, they generate similar sets of displays going from the HP result to close to the HE result. There are however subtle differences between the three techniques which a simulated test pattern image will clarify below. Further, in a hardware real-time embodiment, their noise characteristics should differ (see Section 4 for further discussion).

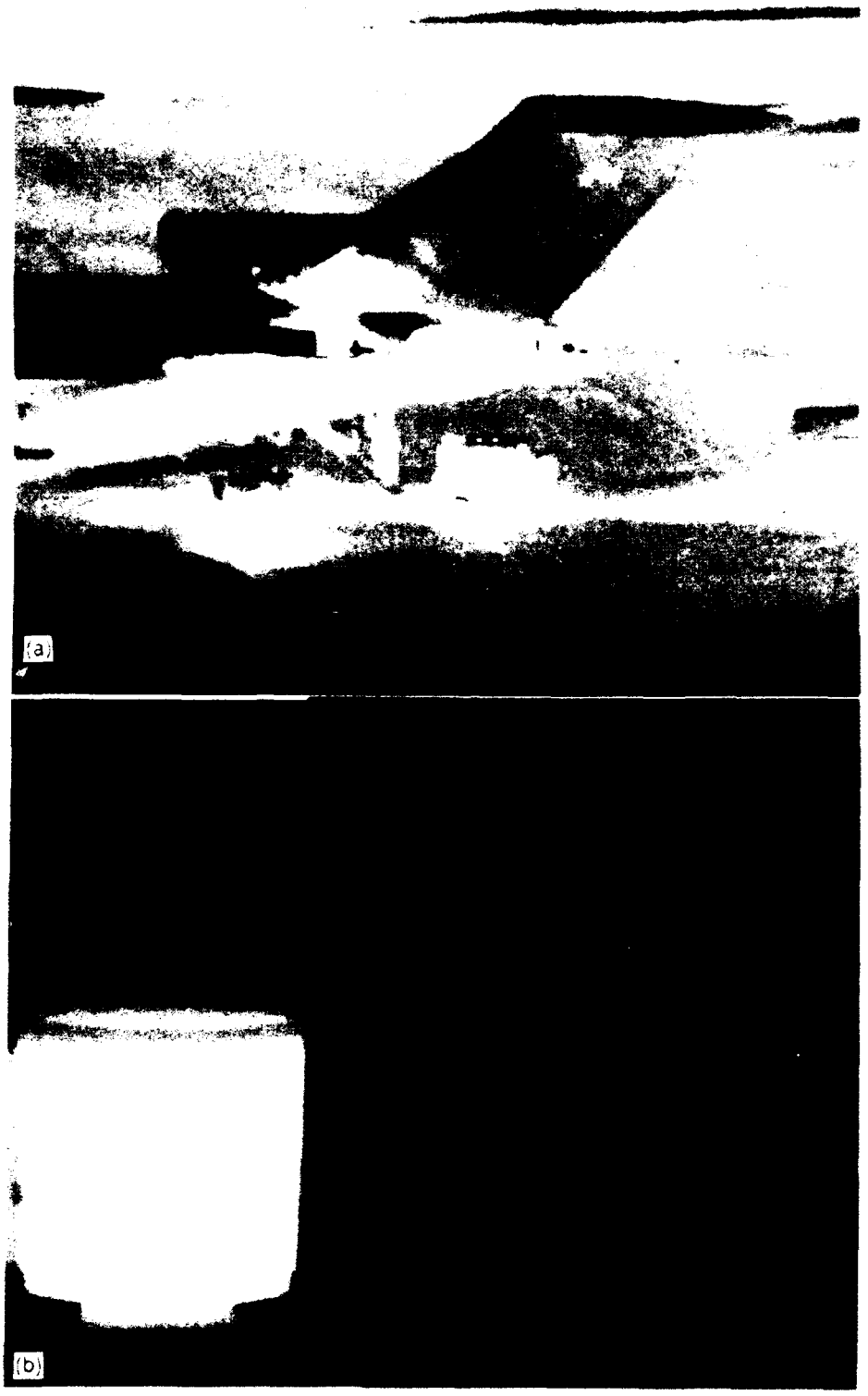



Figure 18. Hybrid displays using Eq. 15.6 with $W = 0.75$: (a) airport; (b) cups.

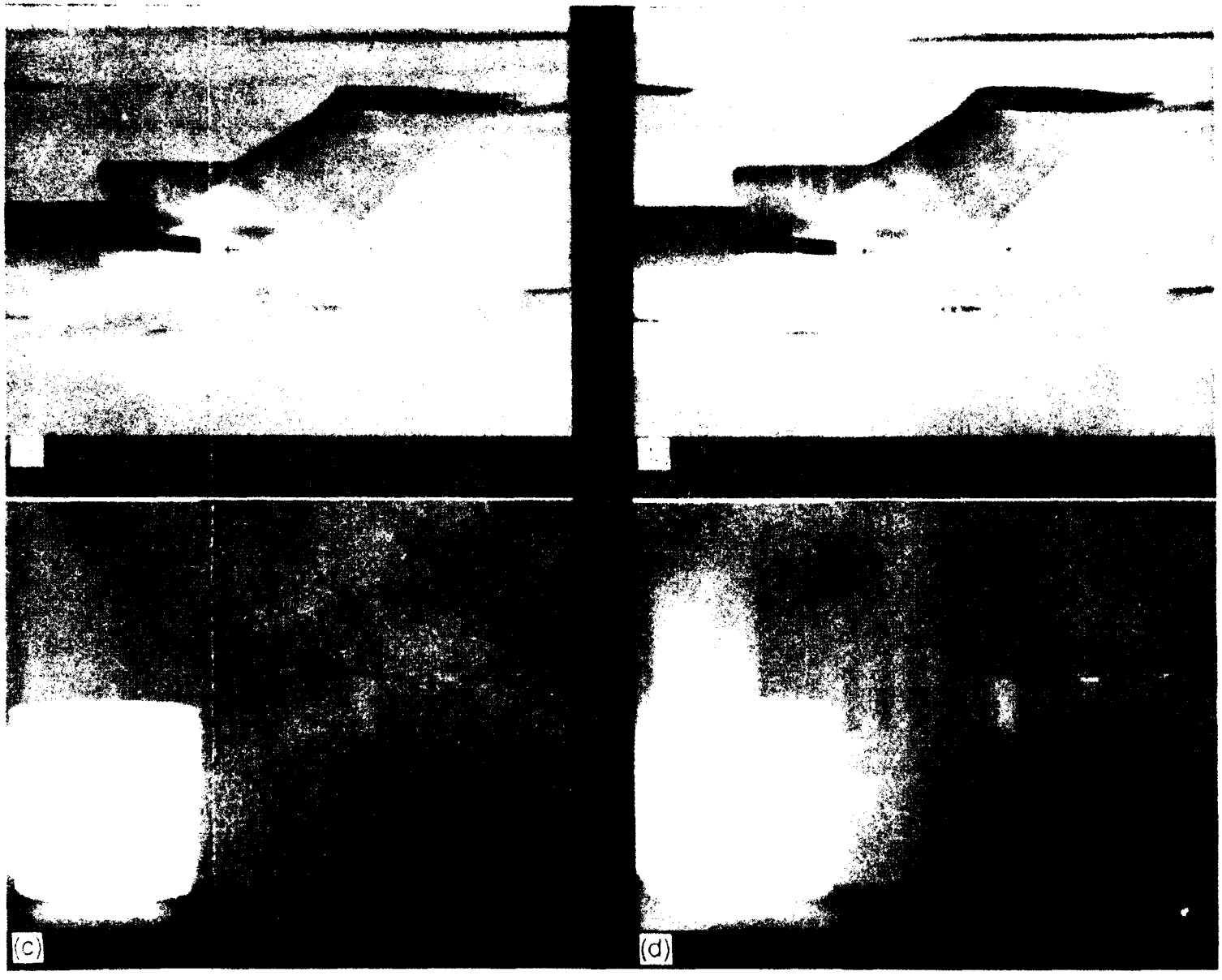



Figure 19. Examples of UP displays: (a) airport at 1/4; (b) airport at 1/8; (c) cups at 1/16; (d) cups at 1/32.

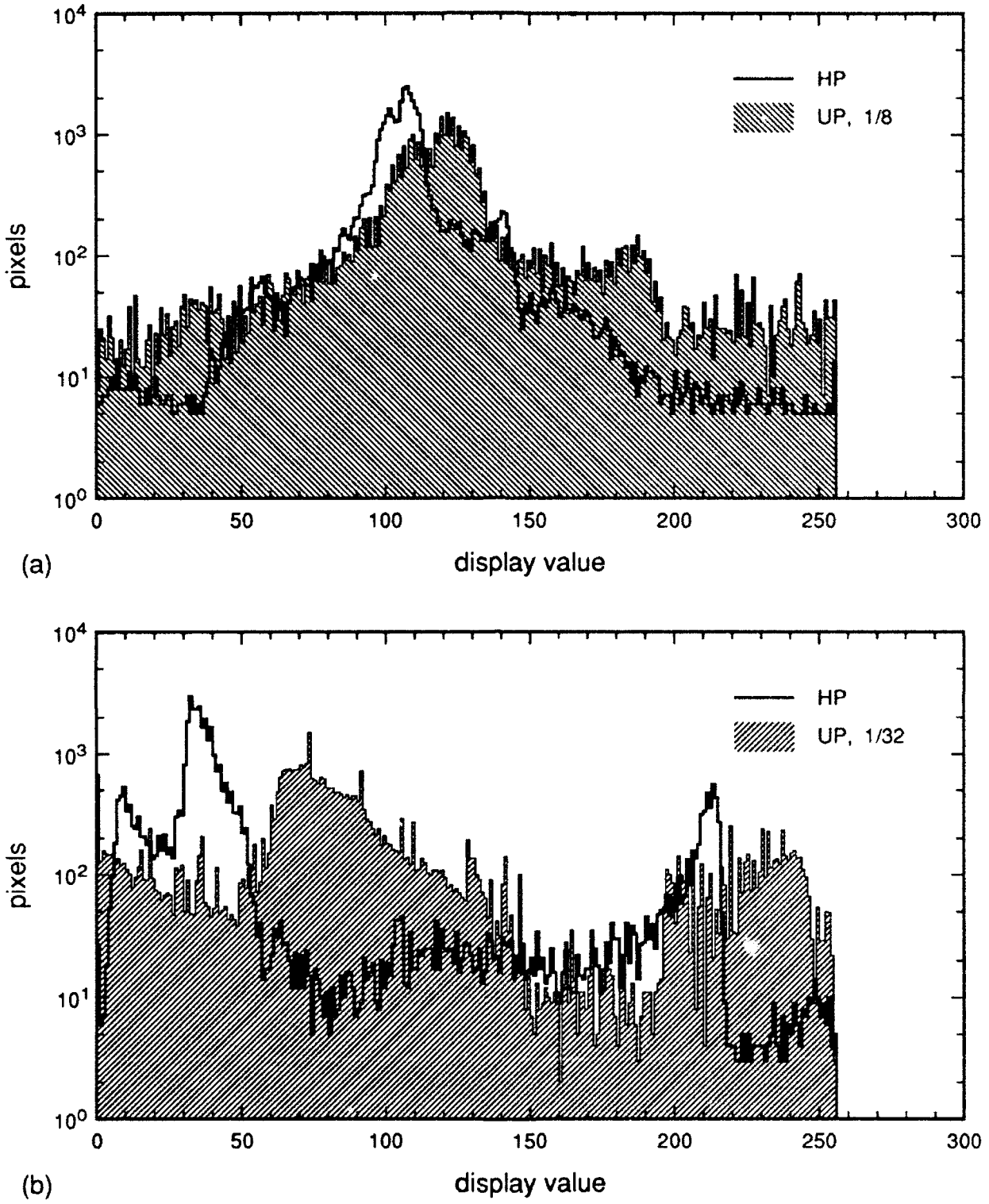



Figure 20. Examples of HP and UP histograms: (a) airport, HP and UP at 1/8; (b) cups, HP and UP at 1/32.

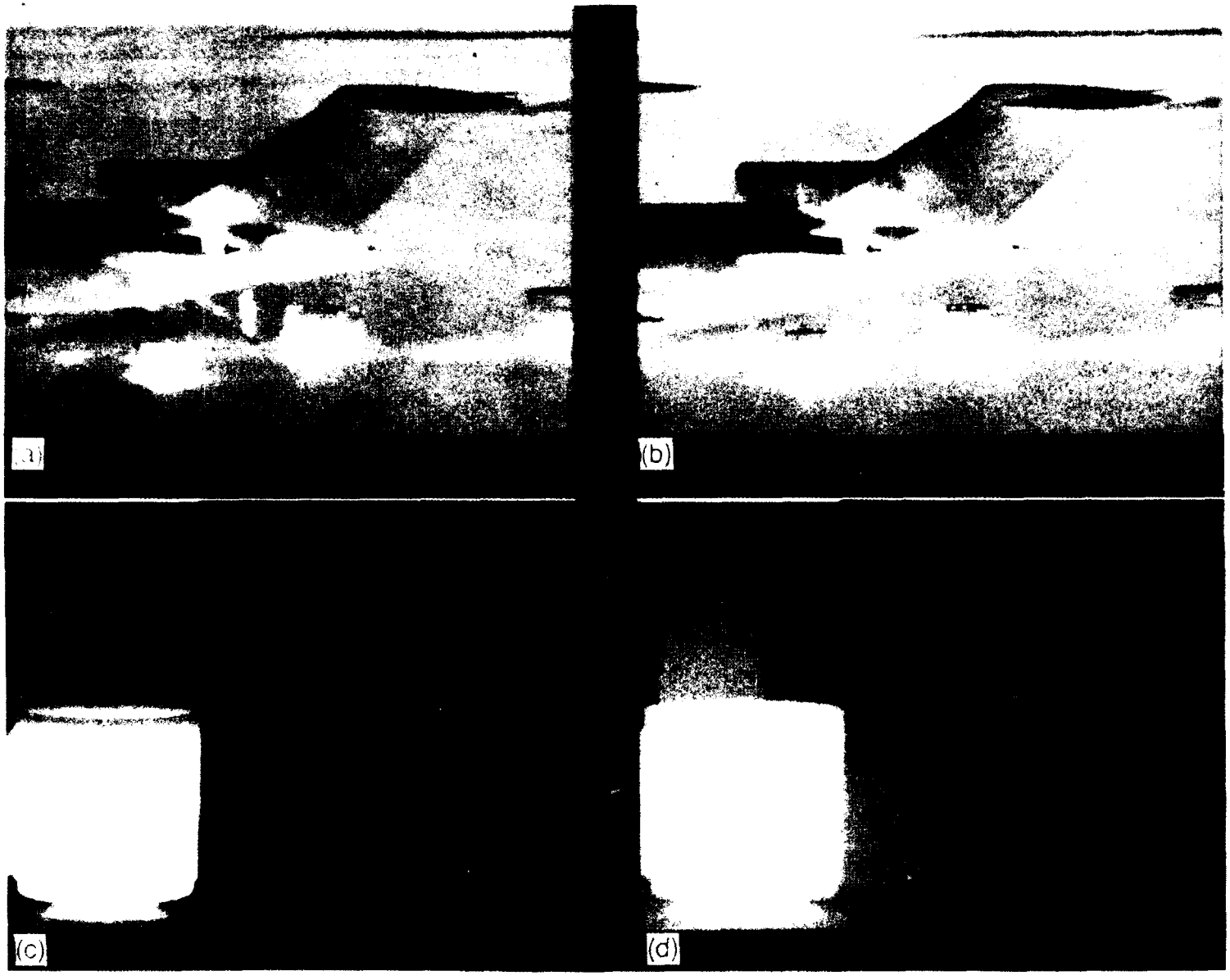



Figure 21. Examples of TP displays with threshold at: (a) 2 pixels; (b) 4; (c) 4; (d) 8.

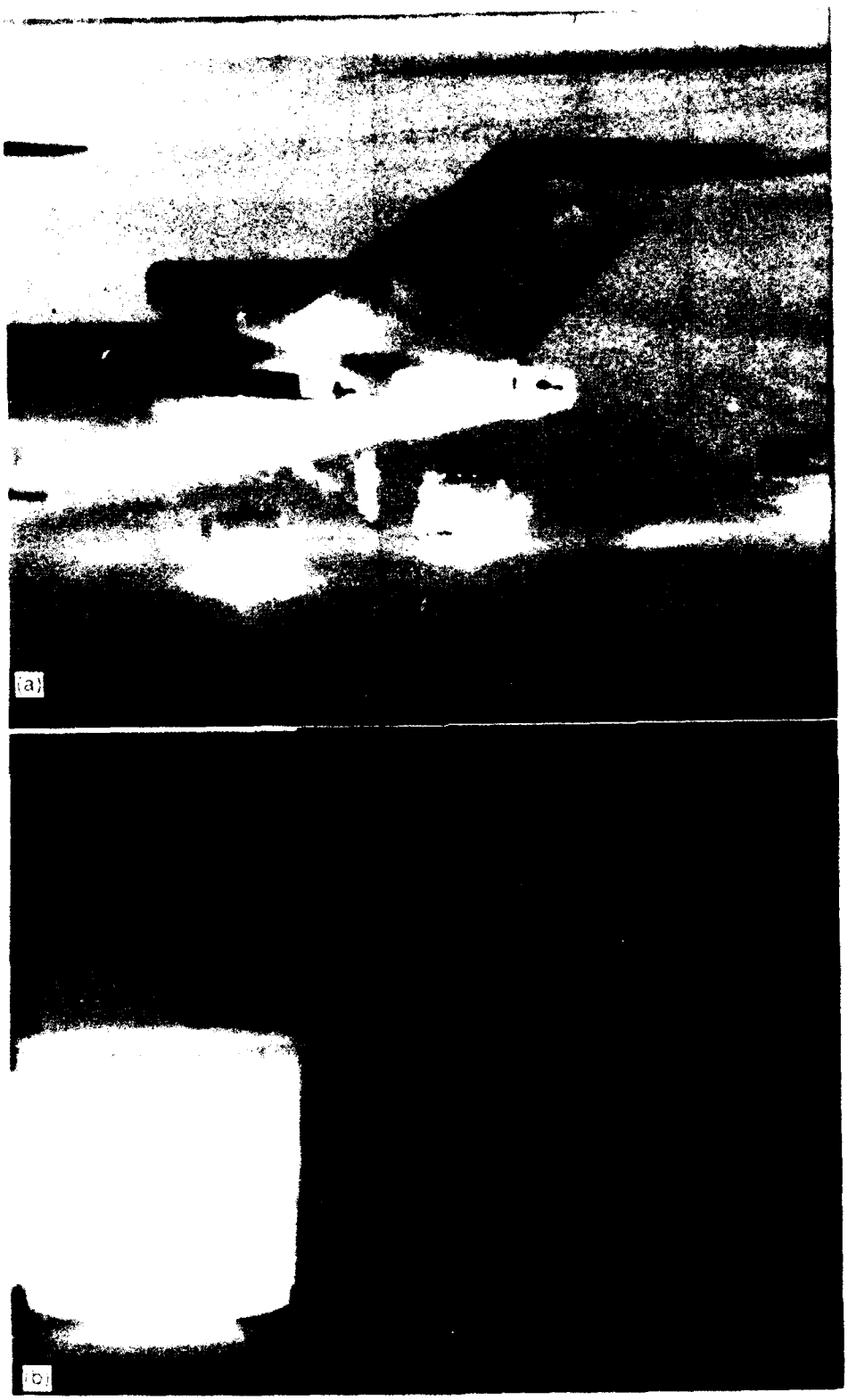



Figure 22. Examples of PE displays with plateau at: (a) 10 pixels; (b) 20.

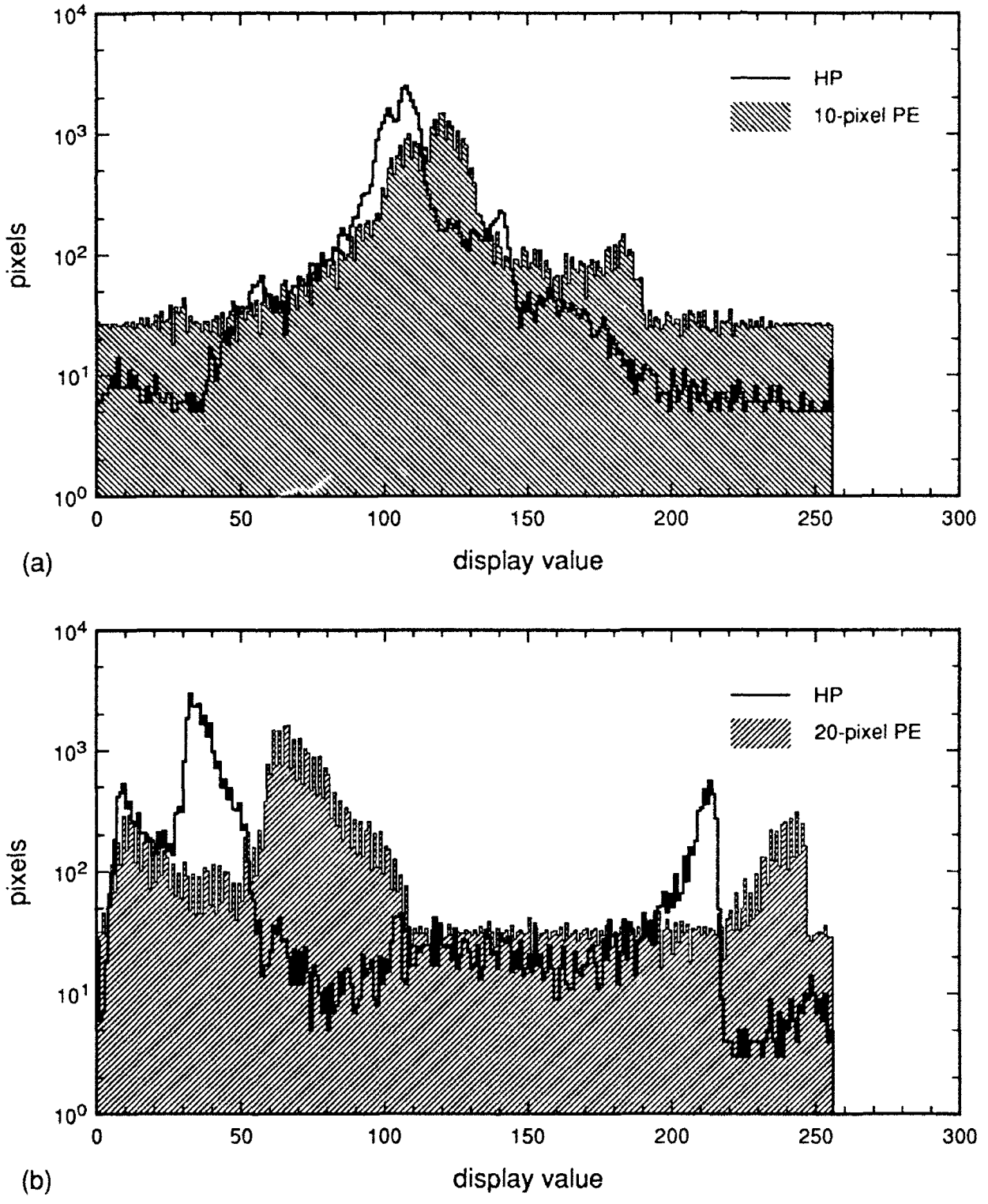



Figure 23. Examples of HP and PE histograms: (a) airport, HP and PE with 10-pixel plateau; (b) cups, HP and PE with 20-pixel-plateau.

Under-sampled projection (UP) is the simplest way to gradually introduce the weighting characteristic of HE – one that in fact further reduces the computational overhead. By calculating a binary histogram based only on every second, fourth, eighth, etc. pixel, one gradually increases the display allocation of more heavily occupied levels. In effect, this is occurring on a probabilistic basis, by not detecting sparsely occupied levels; this leaves more display dynamic range for the occupied ones. The displays for 1/4 and 1/8 under-sampling for the airport scene and for 1/16 and 1/32 for the two-cups image are given in Figure 19. The display histograms for HP and the more undersampled UP are compared in Figure 20. Note the effect on display range allocation of increasing the weighting given to pixels which are at frequently occurring levels.

An alternative to under-sampling is to require a threshold number of occupancies, 2, 4, 8, etc., before deeming a level "occupied" (TP). TP shifts the dynamic range allocation similarly to UP, by preferential detection of more densely-occupied levels, but operates in a deterministic mode rather than a probabilistic one. On actual imagery, very similar results are obtained (Fig. 21; display histograms, not shown, are much like those in Fig. 20).

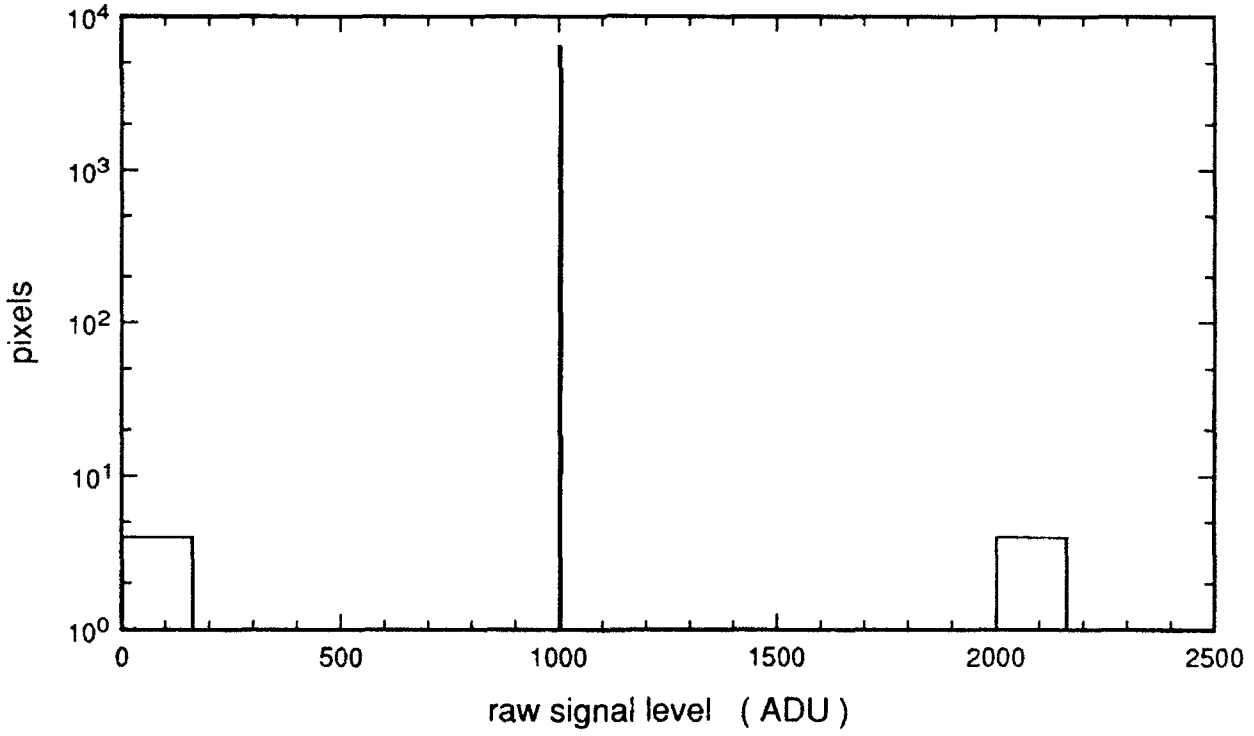



Figure 24. Histogram for test pattern.

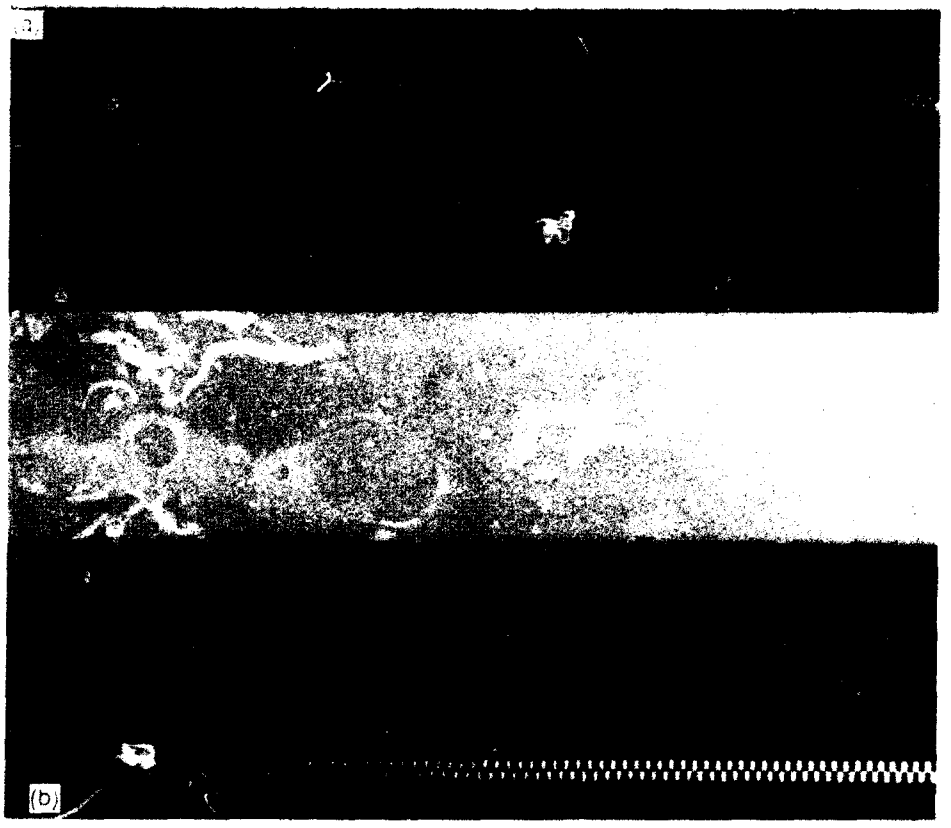
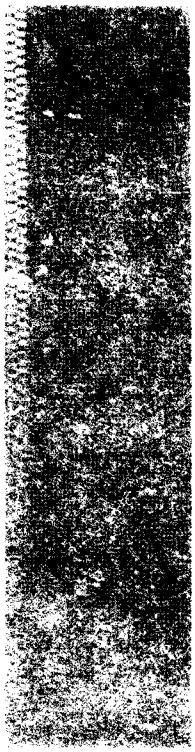
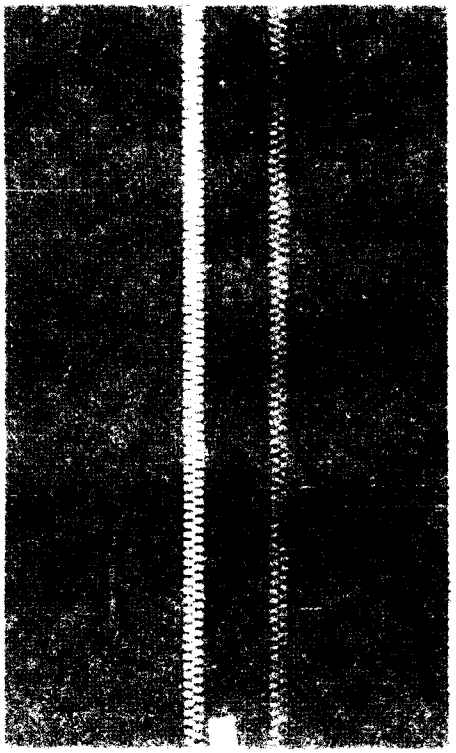
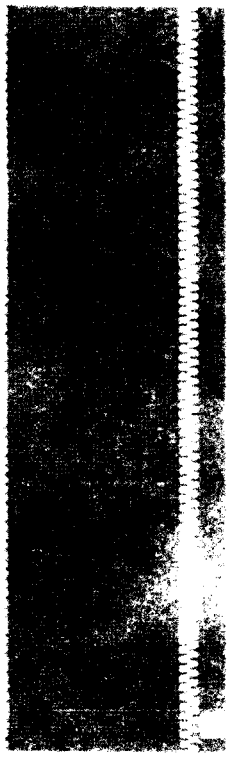
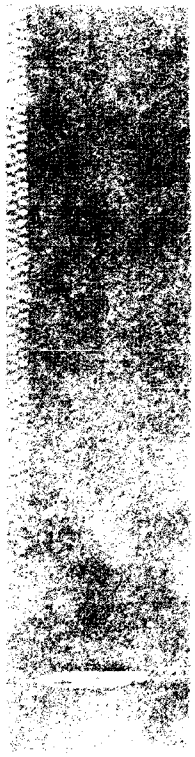


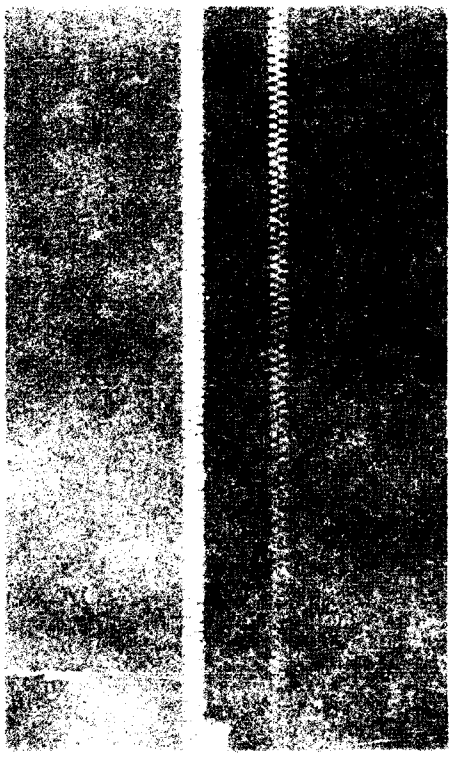

Figure 25. Comparison of HE and HP displays of test pattern: (a) HE; (b) HP.

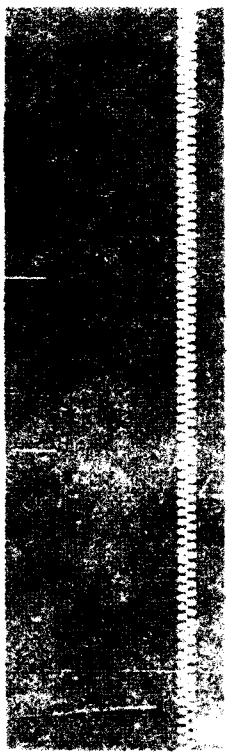












In principle, a better way to blend the benefits of projection and equalization is to perform HE with a cutoff saturation level or plateau imposed on the histogram distribution (PE). Typically, a plateau of 10 to 20 counts is optimum for the 160 × 244 arrays used to generate the imagery here; the desirable plateau would increase with total pixel count. In PE, one generates the $F[i]$ function used in Eq. 3 as before from the full histogram but in effect suspends counting the occupancies for a given level when and if the plateau is reached. All levels at or above this plateau are given equal weight as in projection, while levels below are weighted (in compressed form) as in equalization. Typical results are shown in Figure 22, with Figure 23 comparing the display histograms for HP and PE. Note the close similarity with Figures 19 and 20.

The subtle differences in principle (though usually not in practice on current IR imagery) among UP, TP and PE are best revealed through a simulated test pattern such as one whose raw histogram is given in Figure 24. A two-level (1000 and 1004) low-contrast checkerboard pattern in the center of the image leads to the histogram peak of 3200 near level 1000. The intensity-graded wavy line signals at the top and bottom of the image lead to the square peaks of height 4. These arise from 4 pixels each at levels 1 to 162 (top) and at levels 2001 to 2162 (bottom). In all, there are 327 occupied levels. As could be foreseen, HE brings out the checkerboard pattern but loses the gray-scale resolution of the wavy lines (Fig. 25), while HP optimizes the latter but loses the contrast in the checkerboard. The computationally intensive hybrid procedure (75% projection) and PE (40-pixel plateau) afford optimum displays (Fig. 26a, b), bringing out the checkerboard pattern while largely retaining the gray scale gradations of the wavy lines. UP and TP (Fig. 26c, d) shift dynamic range by not detecting levels that are actually present. In the former, segments of the wavy lines are not graded but lumped into fixed display values; while in the latter, when the minimum threshold condition exceeds 4, the checkerboard pattern is vividly brought out but each wavy line has been reduced to a fixed display value.

We conclude this section with the following observation: Our survey of global monotonic algorithms was initially motivated by a pressing need for a real-time automatic contrast control to replace the manual offset/gain controls on IR cameras, which require frequent adjustment; and the HP algorithm has successfully provided the requisite automated and optimized display (see Section 4).

# 3. LOCAL CONTRAST ENHANCEMENT

## 3.1 Introduction

A familiar experience to any operator of an actual IR camera is that in bringing out local details by adjusting one portion of the image for optimum contrast, one will obliterate other parts of the image. As a software example, Figure 27 highlights the rudder of the airplane image and the warm cup portion of the cup image by the assignment of the raw signal span of these parts of the images to the whole display dynamic range. The purpose of the algorithms considered in this section is to perform this enhancement function automatically and simultaneously for all parts of the image with a minimum generation of artifacts.

In order to accomplish such a locally-enhanced display, one must depart from global, monotonic mappings. Low contrast details such as those on the cups are frequently not discernible in such mappings. In the HP algorithm, for example, the total number of occupied levels $N$ is a measure of the signal dynamic range. $N$ is 451, 1282, and 877 levels for the geese, airport and cup images respectively. When one is mapping on the order of 1000 levels of information into 256 nominal levels of display (about 100 discernible levels of gray), about every four adjacent occupied levels will be fused to the same display value, and raw signal levels 8 apart or greater can end up as indistinguishable adjacent display values, though a representative noise level[5] is only 3 to 5 for the 12-bit raw signal. Clearly, real information can be lost in such global, monotonic mappings. (In on-line imagery, the spatial and temporal averaging performed by the eye[15] would lower the quoted noise levels slightly.)

In the next three subsections we describe three distinct categories of algorithms for locally-enhanced display.

## 3.2 Local Implementation of Global Algorithms

Many locally adaptive enhancement methods described in the literature take advantage of the obvious fact that the dynamic range of a sub-image is typically less than that of the total image. For example, in the HP algorithm, the degree of contrast hinges on the number of occupied levels. To the degree that the local occupancy differs from the total, one can increase the display contrast by a local application. In Figure 28, the results in applying the HP procedure to 4, 8, and 16 disjoint sub-images respectively are shown for the cups image. As the number of sub-image divisions increases, the degree of contrast expansion does also,

[15]Mooney, J. M. (1991). Effect of spatial noise on the minimum resolvable temperature of a staring sensor, Appl. Opt., 30: 3324.

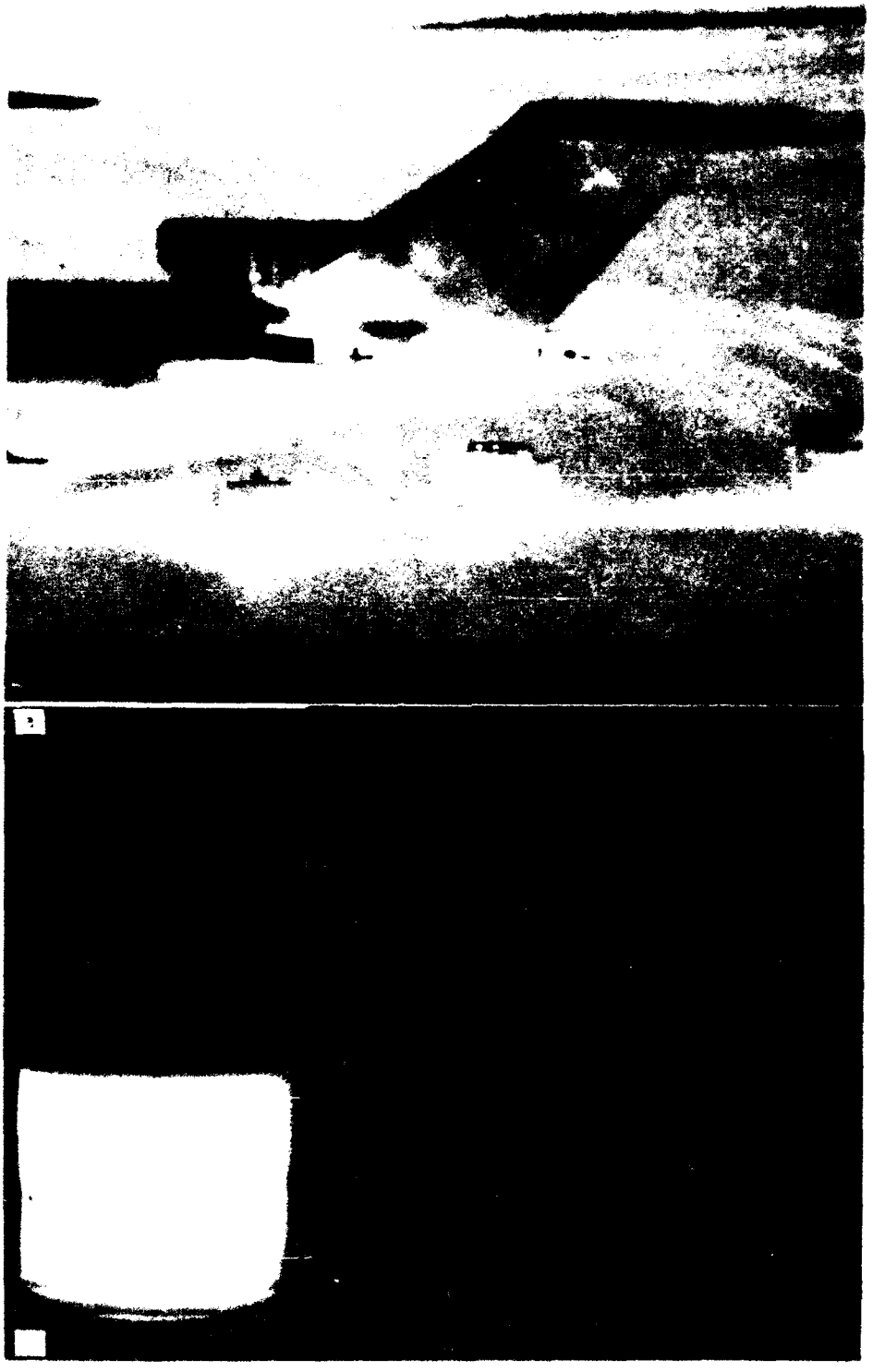


Figure 27. Examples of mapping part of raw sensed range to full 8-bit display range: (a) 1800 – 1900, (b) 2300 – 2360.


a
b

but so does the number of luminance "seams" at sub-image borders; this tends to distract the eye, and has the potential of destroying information. While this is the simplest way to apply a global algorithm locally, better results are achieved (at greater computational cost) by using sliding or overlapping windows or interacting sub-images. Both overlapping- and sliding-window implementations of HP will be described below. Tom and Wolfe[12] describe local implementation of HE with a sliding window.

A clever example of local implementation of a global algorithm of the linear type described in Section 2.2 is local range modification (LRM) of Fahnestock and Schowengerdt[16] (also see Schowengerdt,[17] pp. 66-67). The image is partitioned into disjoint sub-images. The maximum and minimum signal level of each region are determined. One then assigns to the corner of each sub-image the maximum (minimum) of all the maximums (minimums) of the sub-images which share this corner. Despite the disjointness of the sub-images, the allowed interaction of neighboring regions through the assigned corner values is similar in effect to using overlapping windows. For each pixel, a local maximum and minimum are computed from a bilinear interpolation of the corner values and then used to define a linear local contrast stretch. We have implemented versions of this algorithm for IR images with some results given in Figures 29 and 30.

For the airport image, LRM yields a display similar to local overlap projection described shortly (see Fig. 33). Results are shown for block sizes $20 \times 61$ pixels (32 total sub-images) and $20 \times 31$ pixels (64 total sub-images). The false changes of luminance are artifacts which are also found in local overlap projection. A more serious weakness of the LRM procedure is its sensitivity to the influence of outlying or fallacious pixels such as found in the first few rows of the cups image or to very strong edges in the image such as the cup boundaries. Large regions can be "whited" or "blacked" out because of an inappropriate minimum or maximum in the local stretch equation. In Figure 30, two versions of the LRM algorithm are applied with $20 \times 31$ block size. The versions differ in regard to the detection and attempted rectification of the effects of anomalous pixel values. (Details are not important here. We simply emphasize the great difficulty in making the LRM algorithm robust to such severe artifacts over a range of image types.) If one applies these same two LRM variations to a single-frame, one-point-corrected version of the cups image (Fig. 31), different but equally severe artifacts are present. (This alternate cup image, referred to hereafter as the "noisy" cup image, is more representative of the majority of our images in its degree of temporal and spatial noise (see end of Section 2.1) and is used in the remainder of this section in addition

[16]Fahnestock, J. D., and Schowengerdt, R. A. (1983). Spatially variant contrast enhancement using local range modification, Opt. Eng., 22: 378.

[17]Schowengerdt, R. A. (1983). Techniques for Image Processing and Classification in Remote Sensing. Academic Press, New York.

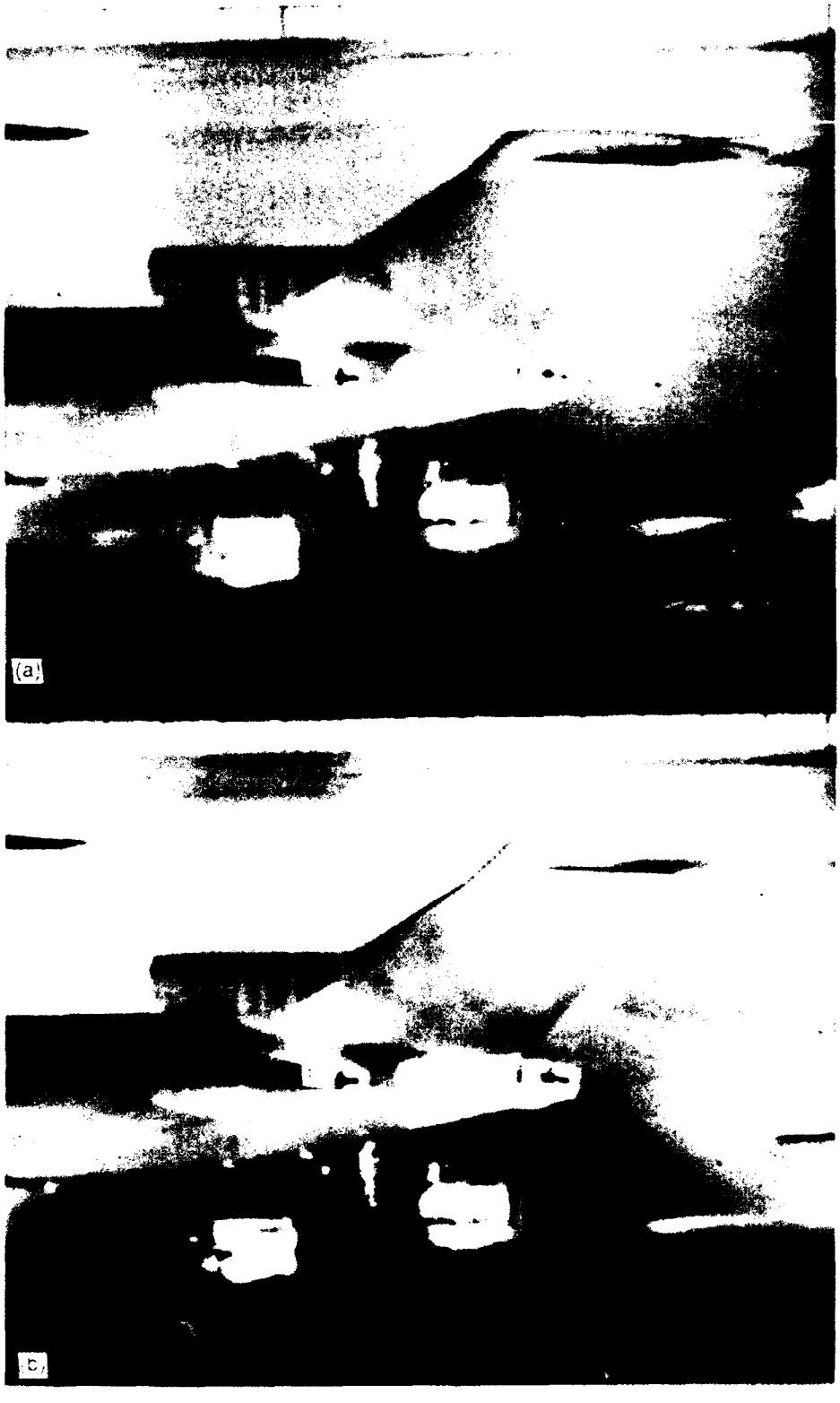
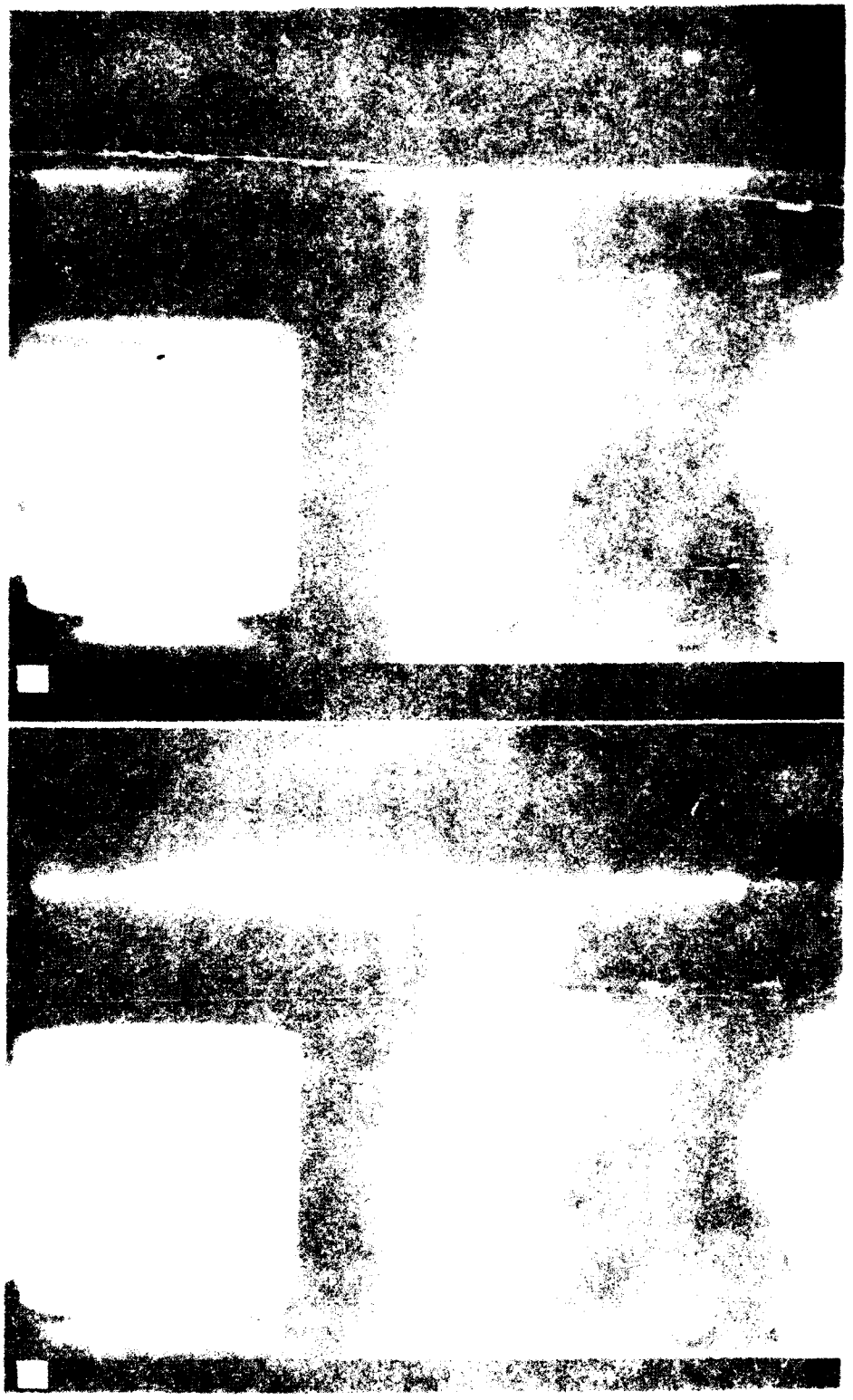



Figure 29. Examples of LRM displays: (a) 20 × 61 block size; (b) 20 × 31.



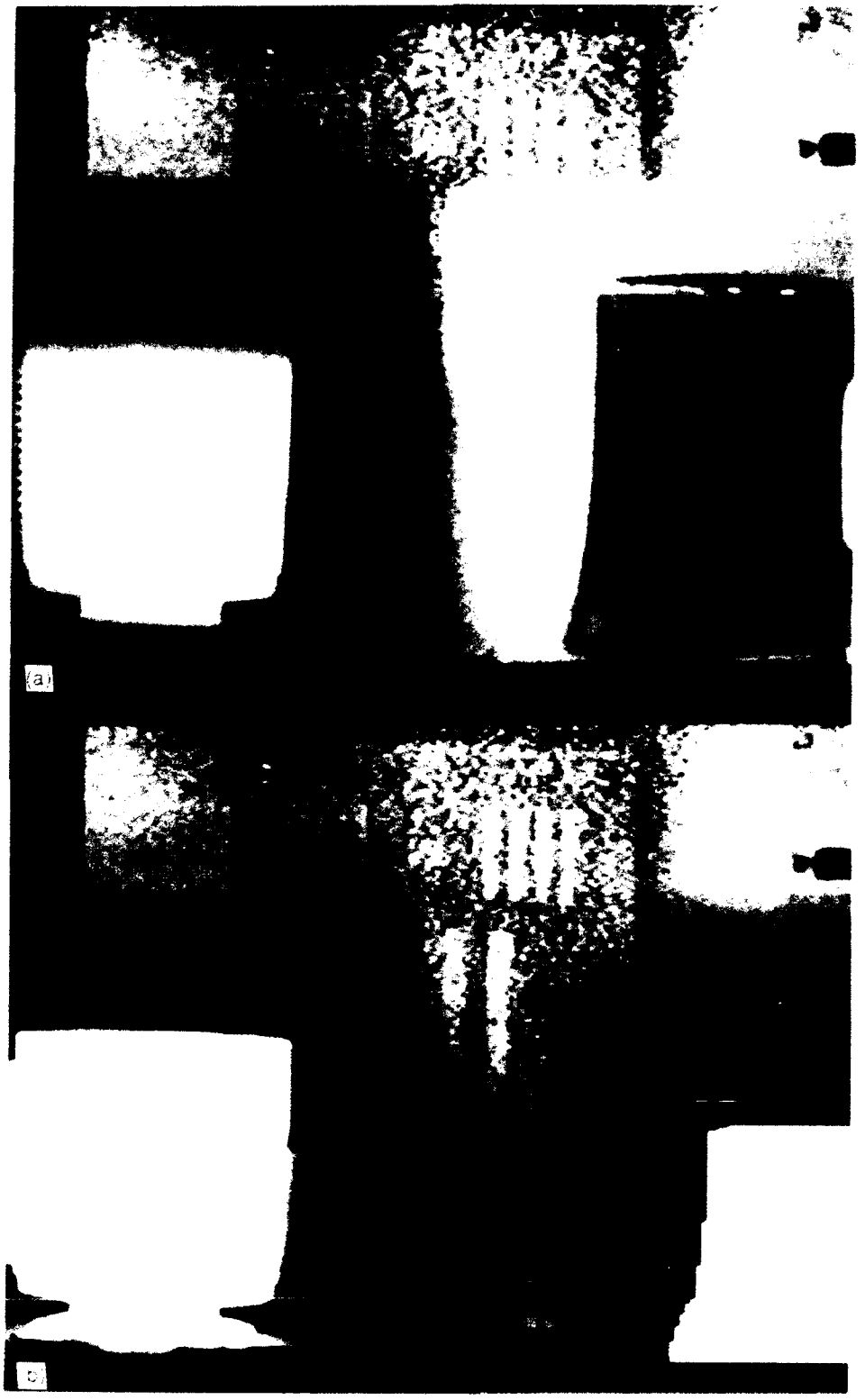



Figure 31. Examples of displays from two versions of LRM, 20 × 31 block size, for noisy cup image. See text.

to the first cups image, in order to illustrate the noise-amplifying effects of local enhancements.)

The LRM algorithm inspired a somewhat analogous but more successful local implementation of a global algorithm, namely a local form of HP based on overlapping windows (overlap projection, OP). OP is implemented after first (conceptually) adding four phantom rows to the image to give 160 columns × 248 rows of pixels. (These pixels have widths twice their height.) The image is then divided into 80 disjoint sub-images (16 × 31 pixels). To eliminate or reduce the luminance seams at sub-image boundaries, one applies the HP transformation within a local window (size 32 × 62 pixels) which encompasses four sub-images (Fig. 32). The HP transformation is performed for each of the 63 distinct positions taken by the window as it translates by half its linear dimension in each direction – thus the overlap. For the four corner sub-images or regions, a unique transformation is defined for each pixel. For the 28

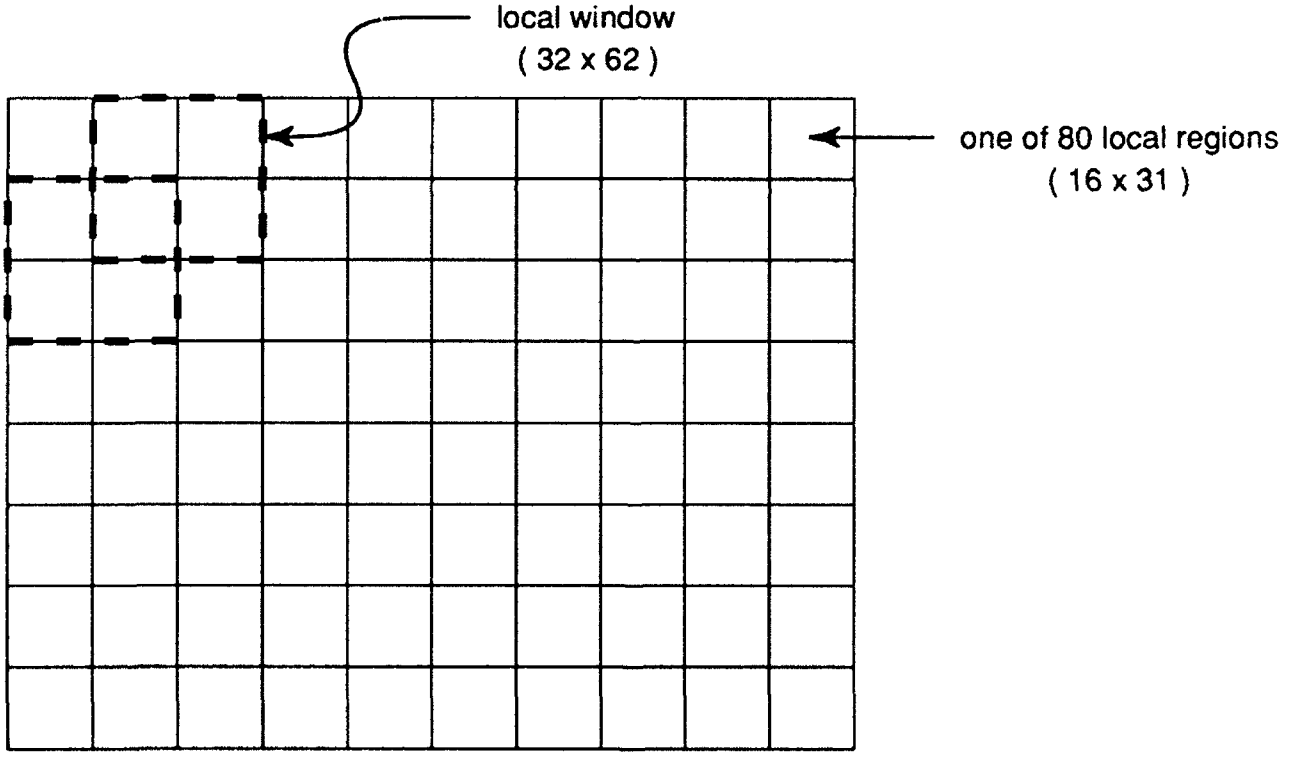



Figure 32. Schematic of the OP technique.

edge regions, two transformations are defined, and for the 48 interior regions four transformations are defined for each pixel. For the pixels in the edge and interior regions, the final display value is generated by linear and bilinear interpolations of the window-defined transformations, respectively.

Results of OP are shown in Figures 33, 34 and 35. Generally, the algorithm is quite successful in bringing out low contrast features such as structural details on the airplane rudder, designs and lettering on the cups, and veins on the hand. False luminance-change artifacts are present, especially noticeable for the cups image. However, rarely do the luminance effects destroy information as radically as in the LRM algorithm. For the hand and Boston skyline images, HP and OP are directly compared. For the latter image, the many levels occupied by the clouds (this image has a total of 2025 out of 4096 possible levels occupied!) compress the global display of the buildings into the low end of the display scale. Hence the local implementation brings out much lost detail. The hand image is a graphic example of the difference between a global-monotonic display and a locally adaptive one.

An alternate means of applying local implementation of HP is by a sliding window approach (SP). Using a window size of $11 \times 11$, $21 \times 21$, $31 \times 31$, or $41 \times 41$ pixels, we compute the HP transformed display for the pixel centered in this window as well as that of the pixel directly one row below. One then slides the window to center the next pair of pixels and repeats the computation. (If the pixels were square, one would do four pixels at a time.) The computationally-intensive SP algorithm is a very strong contrast enhancer (Fig. 36) which brings out much noise as well. It is most useful as a supplement to OP in the smaller window size versions. As its degree of enhancement lessens to approach that of the OP procedure ($31 \times 31$ and $41 \times 41$ sizes), SP has more serious artifacts which are more likely to destroy information.

Two major drawbacks of local implementations of HP are, first, the computational speed and memory requirements of a real-time implementation, and, second, the haphazard interaction between the image and the degree of enhancement afforded by the procedure. The latter reflects the circumstantial dependence of the ratio of local versus global numbers of occupied levels. These drawbacks are circumvented by the algorithms described next.

## 3.3 Modulo Processing

The mappings described in this subsection[2] grew out of the well-known technique of sawtooth scaling, "often used to produce a wide dynamic range image on a small dynamic range display".[10] For 8-bit displays, one reduces the raw signal level modulo 256, i.e., keeps the remainder (0 to 255) after division by 256. The problem with such a sawtooth mapping (Fig. 37; the mapping shown for a raw signal range of 10 bits can of course be continued to

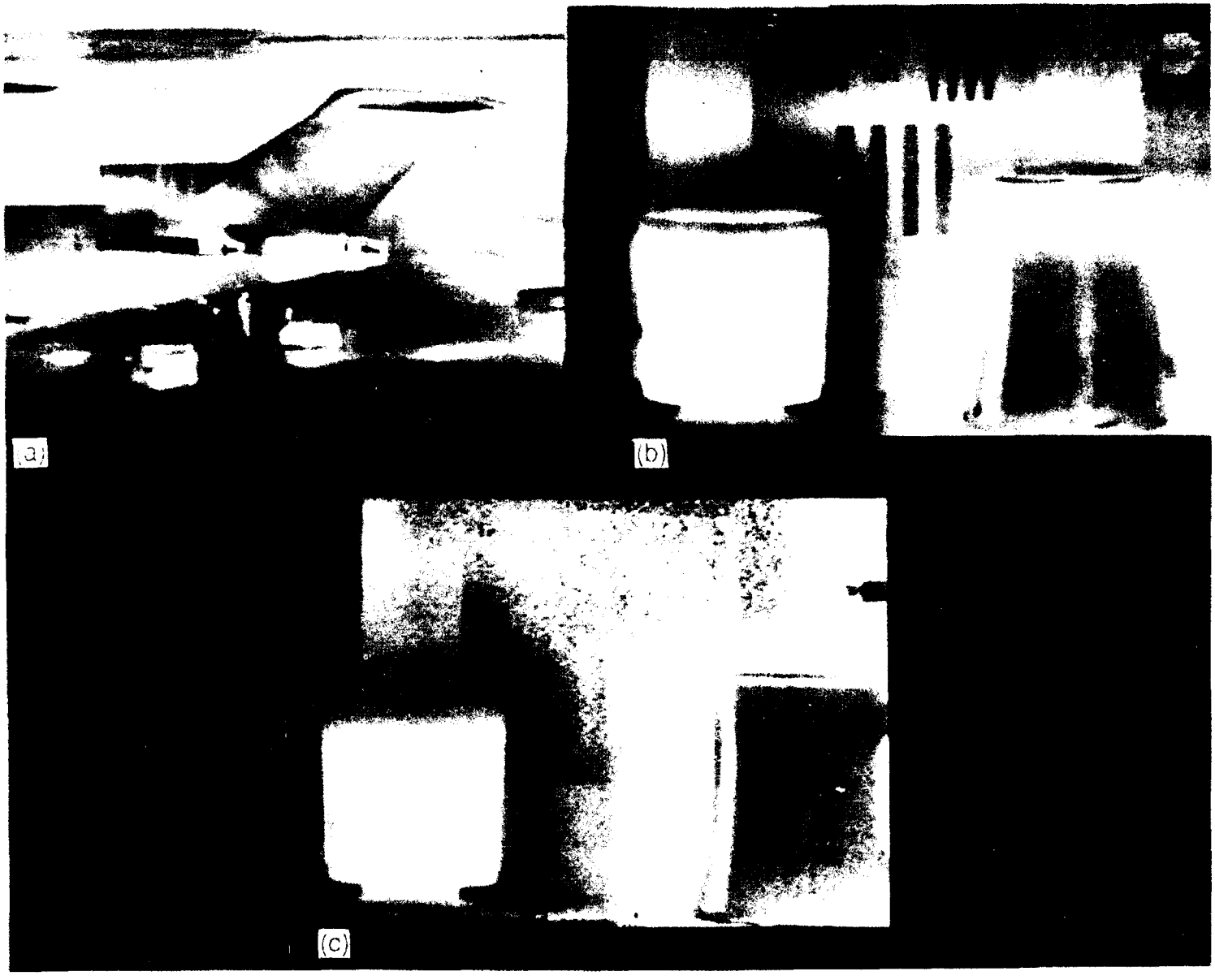



Figure 33. Examples of applying OP to three standard images: (a) airport; (b) cups; (c) noisy cups.

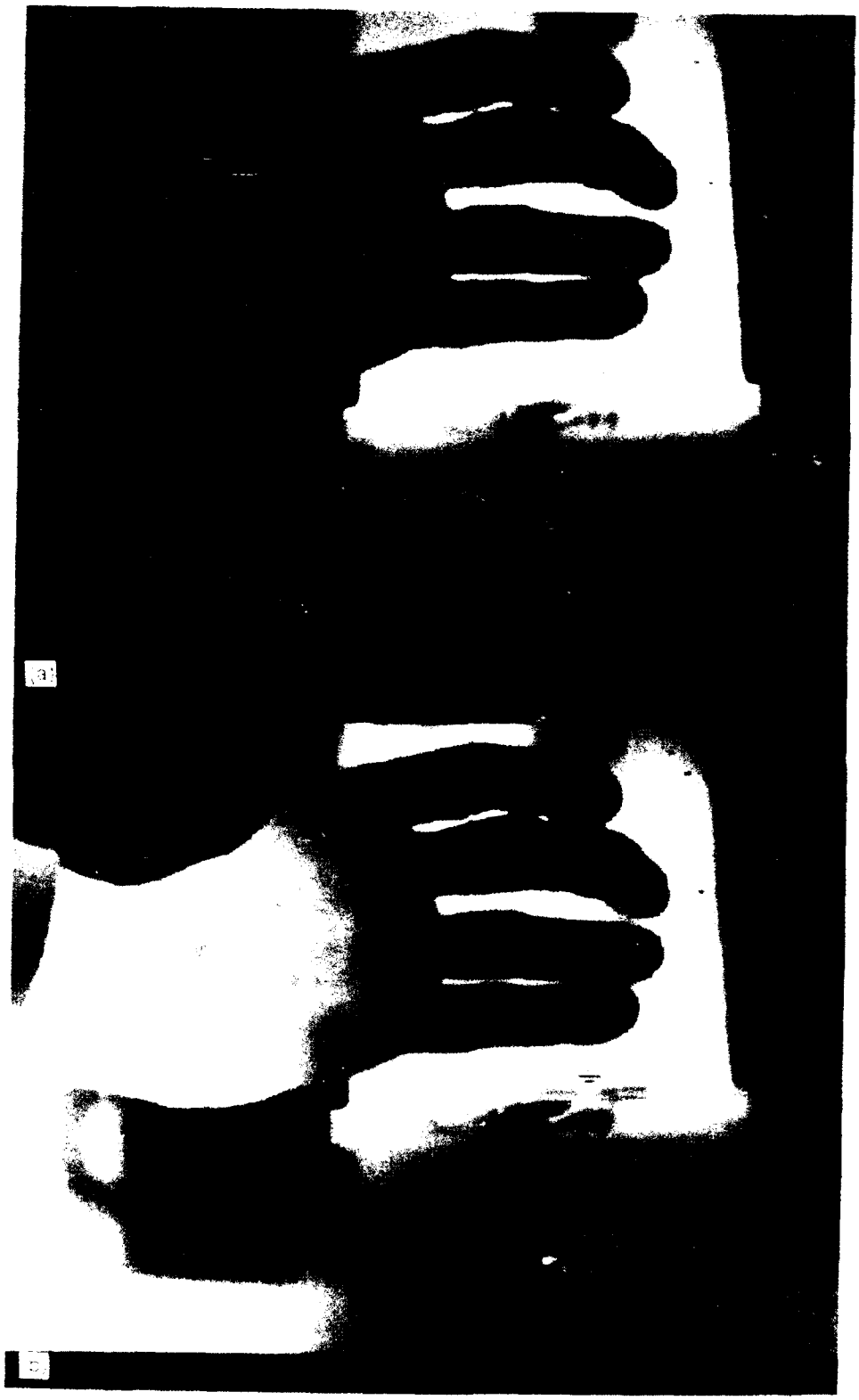



Figure 34. Comparison of HP and OP: (a) HP; (b) OP.

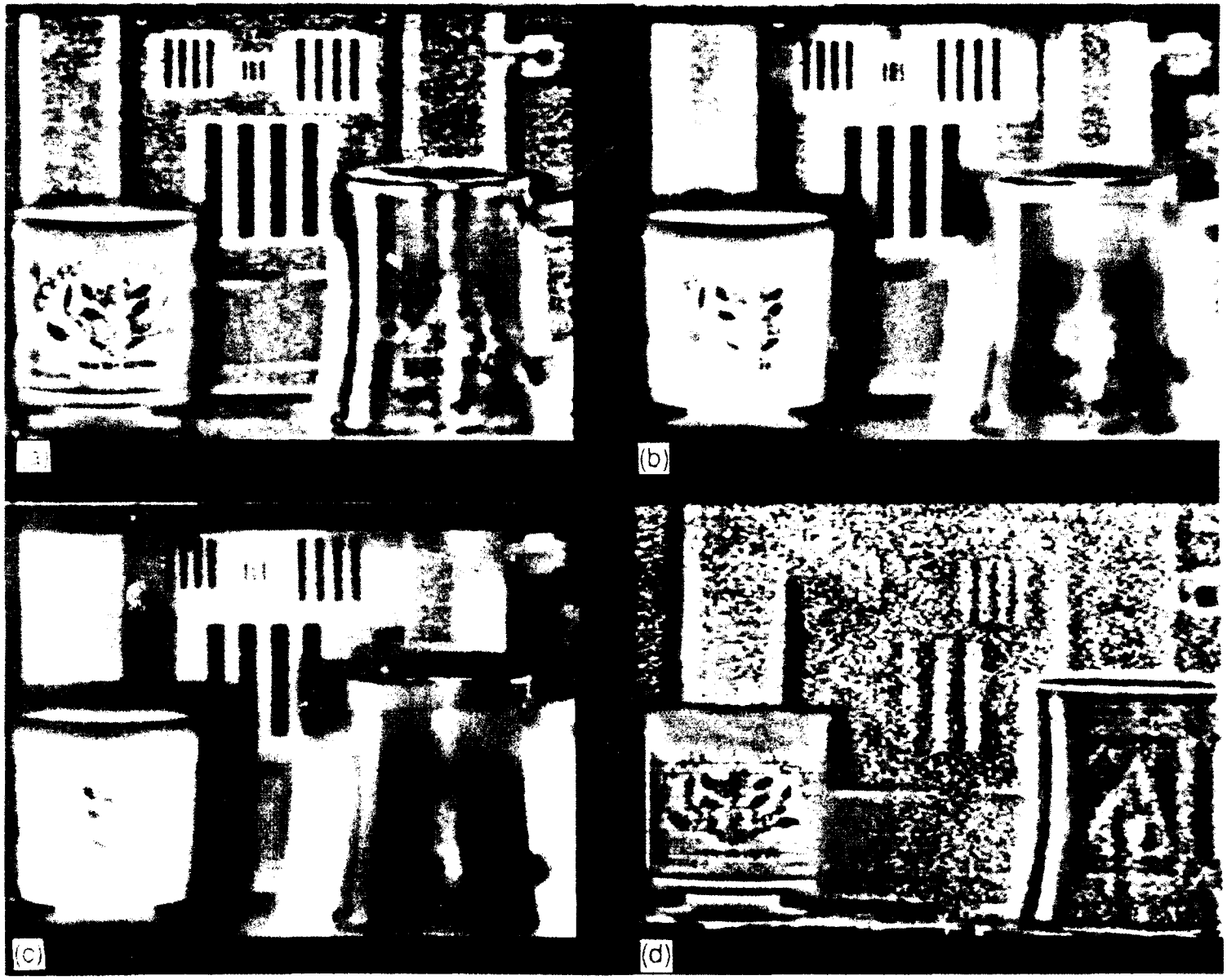



Figure 36. Examples of SP displays of cup images using various window sizes: (a) 11 × 11; (b) 21 × 21; (c) 31 × 31; (d) 11 × 11 (noisy cups).

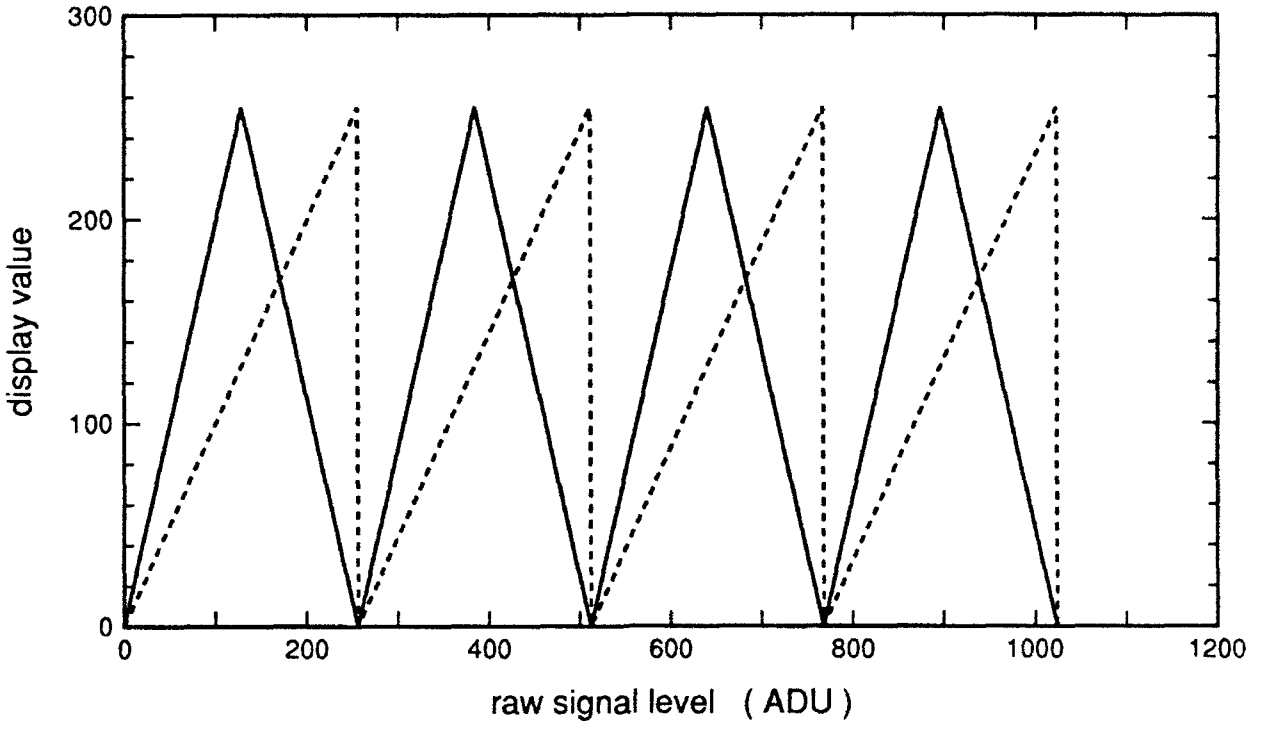



Figure 37. Sawtooth (dotted line) and modulo (solid line) mappings.

arbitrary size) is the discontinuities at multiples of 256 in the raw signal, which lead to sudden black/white or white/black transitions in the display. The simple modification (labeled as "modulo" in the figure, although strictly speaking the sawtooth mapping is the mathematical definition of modulo) to a continuous triangular mapping with the same periodicity increases the utility of the procedure. The asymmetry of the sawtooth mapping is replaced by the symmetry of the modulo mapping (mirror planes at multiples of 128 in the raw signal). In terms of the modulo-reduced value $m$, which ranges from 0 to 255, the revised mapping is given by

$$\text{display value} = \begin{cases} 2m & \text{for } m < 128, \\ 2(256 - m) - 1 & \text{for } m \geq 128. \end{cases} \tag{7}$$

Eq. 7 generates the displays of our (now four) standard images shown in Figure 38. This algorithm is referred to as raw modulo (RM), as the modulo property takes us close to the raw signal data itself. As an image independent global many-to-one mapping (in contrast to the other algorithms in the previous and next subsections), it can be simply implemented in table look-up format, while affording a strong and relatively artifact-free form of local contrast enhancement.

The main drawback of the RM technique occurs whenever a local low contrast feature straddles a mirror plane in the mapping (cf. the lettering in the lower left-hand corner of the original cup image in Fig. 38) where the deviation from monotonicity garbles the display, or whenever regions with high spatial frequency information map into one or more periods in the display, leading to a too rapidly changing, confusing display (cf. the vehicle grilles in the airport scene). Two simple adjuncts to RM can address these problems and increase the utility of the procedure. The first fix is to allow division of the raw signal data by factors of 2, 4, etc. before performing the modulo reduction – this is tantamount to increasing the period of the mappings in Figure 37 by the same factors. The result of division by 2 and 4 for the airport scene (Fig. 39) provides more "readable" displays. A sequence of RM displays "toned down" by increasing factors of two, as in the face image (Fig. 40), often provides a useful survey of the information content of an image. The mirror-plane artifact stems from the arbitrary "phase" in the raw signal (additive constant) with respect to the modulo reduction. The simple fix is to shift the raw signal data by values such as ±32, 64, 128 before the modulo step (Fig. 41). Note the improved reading of the lettering in the lower left-hand of the warm cup in the mod -32 version and the clearer view of the plant logo and "coffee connection" letters in the mod +128 version. However, a more powerful version of this fix can achieve the effect of shifting the raw signal data by any integer value at very little computational cost. We merely employ the sawtooth mapping but display the 8-bit results with a cyclic gray scale: for example, a scale with front-to-back replications of the original, black (0) ... white (128) ... black (255). Then the effect of any data shift can be obtained practically instantly merely by a cyclic shift in the "colormap" for the monitor.

The RM algorithm involves no histogram processing, is simple and effective, but ignores the dynamic range requirements of the particular image. A modulo projection technique (MP) is a more elaborate algorithm designed to adapt to the image dynamic range as measured by the total number of occupied signal levels $N$. The following version was "tuned" to the present IR images – in principle it can be adjusted for other kinds of imagery. If $N$ is less than 512, one applies the RM algorithm except that for additional contrast enhancement, the raw signal data is doubled (if $N < 100$) or multiplied by 1.5 (if $100 \leq N < 512$) before the modulo 256 reduction. For $N \geq 512$, the occupied levels are ordered from 1 to $N$ (as in the HP technique) and $m$ in Eq. 7 is taken as the modulo-reduced occupancy number rather than the modulo-reduced signal level. For still higher dynamic range images ($N \geq 800$), neighboring occupied levels are coalesced to some extent – but by a factor of two less than occurs automatically in using the HP algorithm – before the modulo 256 reduction.

Figures 42 and 43 show enhancements produced by the MP algorithm. As the images in these figures have $N > 512$, MP acts largely as a toned-down version of RM. Comparing the airport scenes in Figures 38 and 43, the MP version is rather similar but superior to the RM divided by 4 (Fig. 39) display, with sharper structural details on the plane rudder. With very low-noise imagery such as the original cups image, the strongly contrast-enhancing effect of

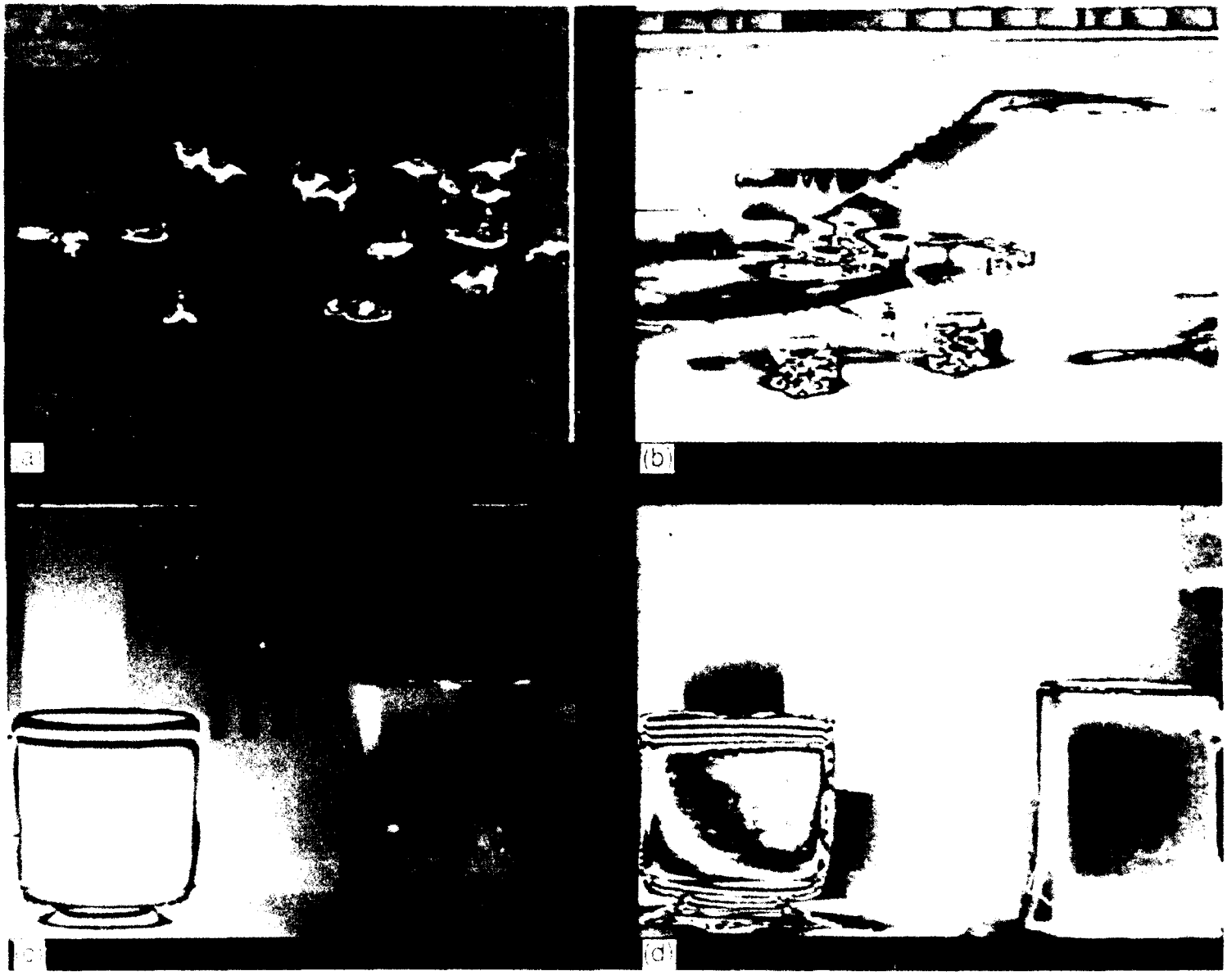



Figure 38. Examples of RM displays of the four standard images.

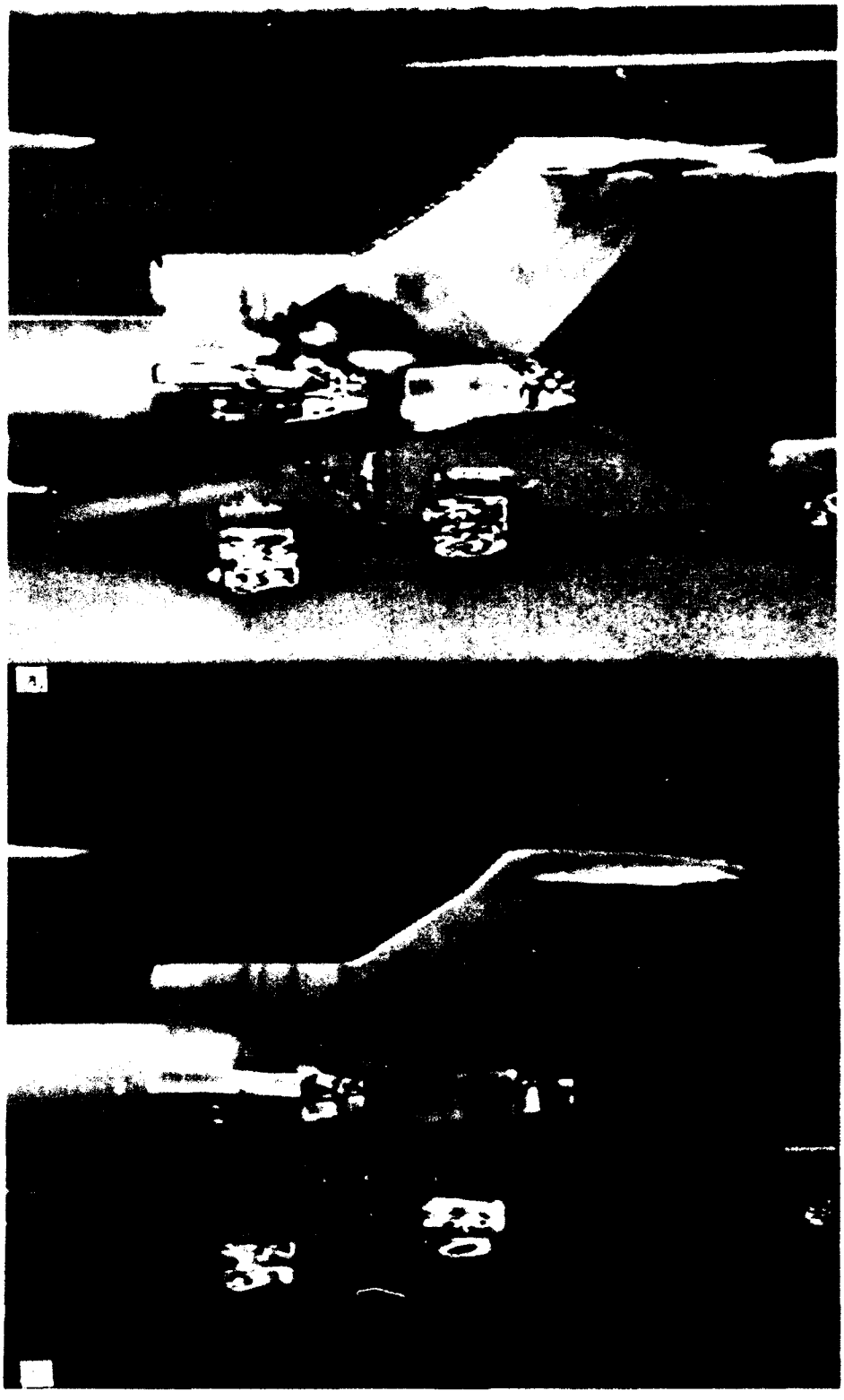


Figure 39. Examples of RM with raw signal divided by: (a) 2; (b) 4.

Figure 36. Examples of RM with row-wise and diagonal [illegible] for 2-ary 4


Figure 41. Examples of RM with raw signal shifted by: (a) 0; (b) -32; (c) +128.

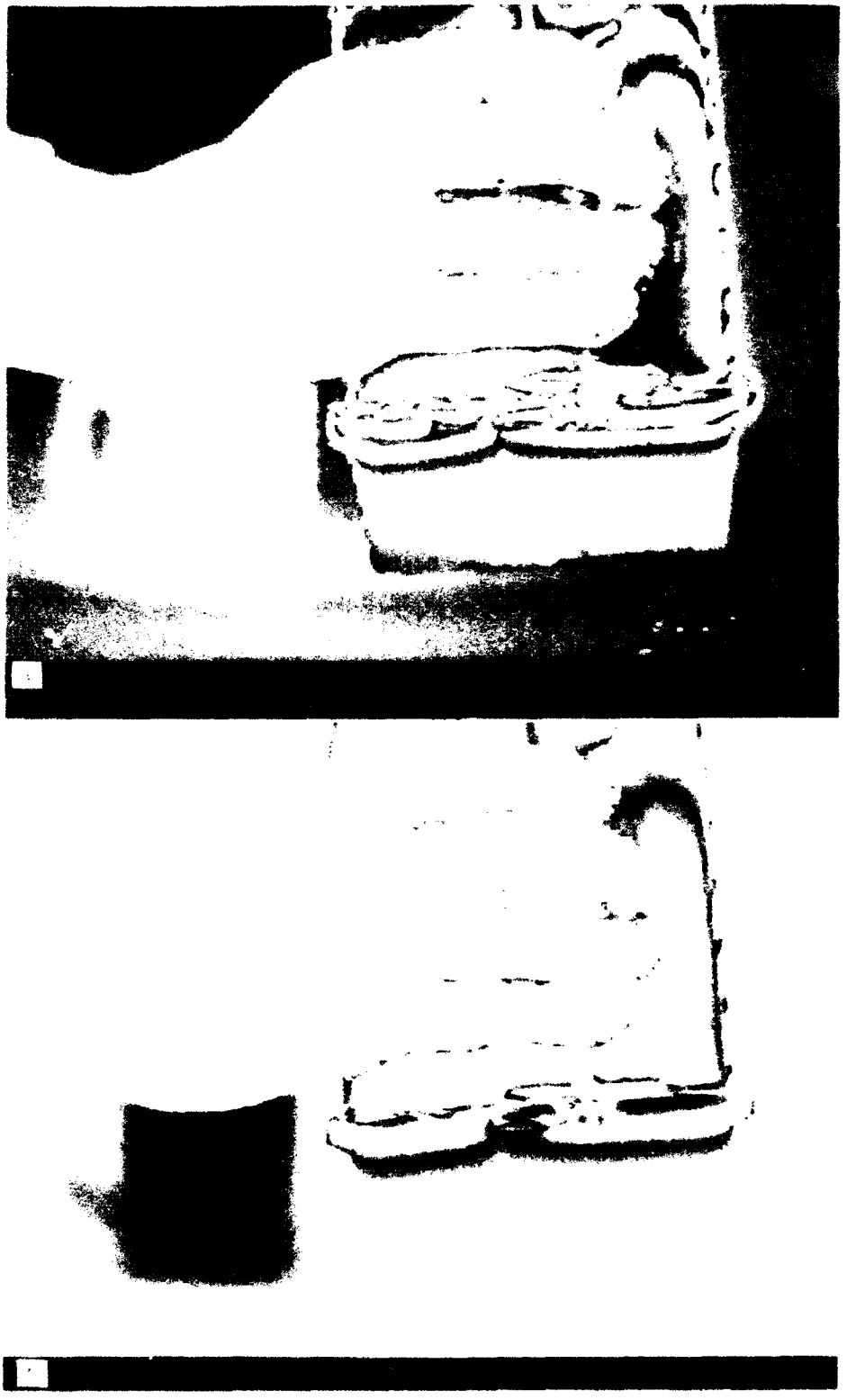


Figure 12. Comparison of RM and MP: (a) RM, (b) MP


Figure 43. Examples of MP displays: (a) airport; (b) cups; (c) noisy cups.

RM works well. For the noisy cups scene, MP gives a clearer view of the warm cup, while RM reveals more of the bar patterns and cold cup details, along with more temporal noise. For the hand image (Fig. 42; compare Fig. 34), both modulo algorithms bring out the veins and sleeve cuff details (note the reverse contrast change in the veins in Fig. 42), with better detail in the fingers than does the OP algorithm. The OP display does retain a better sense of the global thermal sense of the raw signal – this is typically true. More comparisons among the algorithms for local contrast enhancement will be given in Section 3.5.

## 3.4 High Frequency Enhancement

Up to now, we have concentrated exclusively on the spatial-domain point of view (SD). The Fourier domain is now widely used in the analysis and filtering of multidimensional signals[18] because of the wide availability of the FFT routine for computing the discrete Fourier transform (DFT). For images (2-D signals), the spatial frequency Fourier domain (SFD) viewpoint is widely employed for the "restoration" problem in image processing but in general only lip service is paid to the SFD in the "enhancement" problem in image processing. (An exception is the excellent book by Wahl[19] in which the DFTs of images are used throughout to underscore trends and basic principles.)

The algorithms treated in this subsection, high frequency enhancement with linear filters, can be implemented in either the SD or SFD domains. To our knowledge, there are only two standard algorithms designed strictly for SFD implementation. One is homomorphic filtering (Wahl,[19] pp. 84-86), in which the signal is modeled as the product of a high frequency reflectance component and a low frequency luminance component. One converts the product to a sum through the log function, enhances the (now) additive high frequency component in the SFD, and exponentiates the inverse-transformed result to recover the processed image. We have tried this algorithm on our IR images without success – not surprisingly, as the underlying model is not suitable for IR images. The second SFD-specific procedure is "alpha rooting" (pp. 124-126 of reference 18) in which one takes the alpha root (alpha < 1) of the magnitude of the DFT but retains the phase. Again, we found poor results in applying this technique to IR images.

Since any algorithm for increasing local contrast more or less involves enhancing higher spatial frequencies at the expense of lower ones, it seems natural to examine this problem in the SFD. Figure 44 shows the DFTs of the airport image after display with the HP, OP, RM,

[18]Dudgeon, D. E., and Mersereau, R. M. (1984). Multidimensional Digital Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey.

[19]Wahl, F. M. (1987). Digital Image Signal Processing, Artech House, Norwood, Massachusetts.

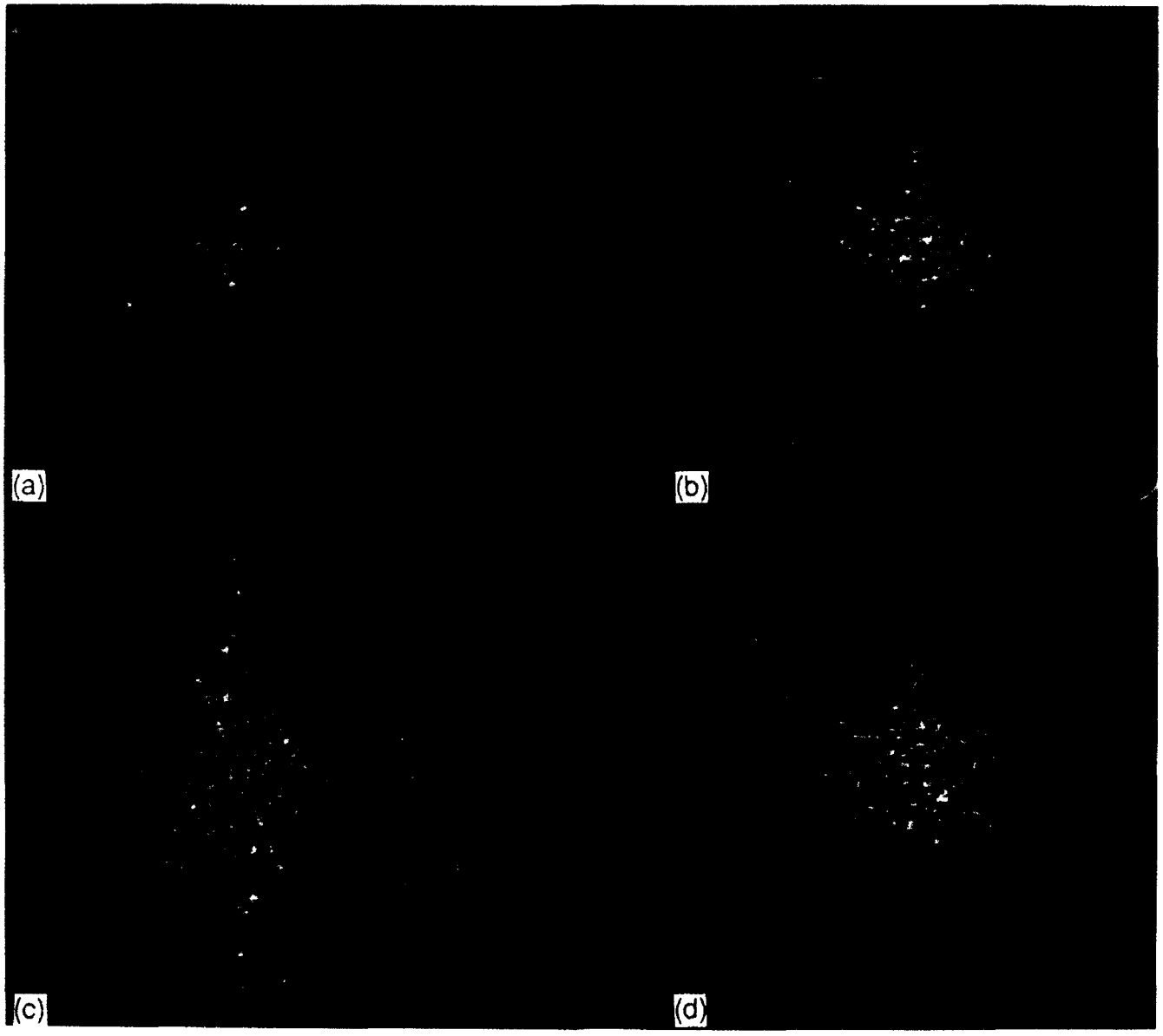



Figure 44. DFTs of the airport scene after applying (a) HP; (b) OP; (c) RM; (d) MP.

and MP algorithms respectively. (We are using 128 × 256 point FFTs on our images by truncating the first and last 16 columns and mirror-expanding the last 12 rows of our 160 × 244 images. More specifically, Fig. 44 shows log-like displays of the squared magnitude of the DFT with the center of symmetry indicating the (0, 0) spatial frequency. See reference 19 for further details.) While the three local algorithms have the general effect of enhancing high spatial frequencies, they do so in a nonlinear and spatially adaptive fashion which would have no counterpart in the SFD. For example, the degree of high frequency enhancement with the OP algorithm varies locally with the ratio of local to global number of occupied levels – a signal dependent and operator uncontrolled process. One would like to accomplish such high frequency enhancement in a controlled and directed manner.

The goal of many high frequency enhancements is to improve subjective image quality from the standpoint of psychophysics by accenting edges, called "edge crispening" (pp. 322-326 of reference 10) or (the term we prefer) "image sharpening". A typical convolution mask for this purpose is

$$1/7 \begin{bmatrix} -1 & -2 & -1 \\ -2 & 19 & -2 \\ -1 & -2 & -1 \end{bmatrix}. \tag{8}$$

Sharpening the raw signal data of the airport image with this mask and then mapping into 8 bits with HP gives the display of Figure 45a. The thermal span (monotonicity) of the image is largely intact but only slight contrast enhancement results. One can implement a graded set of similar operations by means of the equation,

$$P_c' = P_c + a(P_c - \bar{P}_{n \times n}), \tag{9}$$

where $P_c$ and $P_c'$ are the initial and final raw signal values respectively of each pixel centered in an $n \times n$ square neighborhood ($n$ odd). $a$ is a small positive integer which controls the degree to which the difference between the central pixel and the $n \times n$ neighborhood average $\bar{P}_{n \times n}$ is amplified. The choice of $a = 2$, $n = 3$ (Fig. 45b) gives virtually the same result as the mask in Eq. 8, while a choice such as $a = 4$, $n = 9$ (Fig. 45c, hereafter referred to as the WS algorithm for weak sinc sharpened), although more blurry, is beginning to exhibit the degree of local contrast enhancement sought.

We arrived at further improved masks for local contrast enhancement by "tuning" a start from Eq. 9 in the SFD and implementing the result in the SD. Applying the WS algorithm to an "impulse" image and computing the DFT of the sharpened result (Fig. 46a), one obtains the transfer function of this filter. As expected, it is a 2-D sinc function aligned along the axial directions with the requisite number of side lobes from the 9 × 9 neighborhood. We next converted this function into an equivalently strong, circularly symmetric Gaussian (exponential) filter by using the three-parameter form suggested by Wahl (reference 19, p. 85),

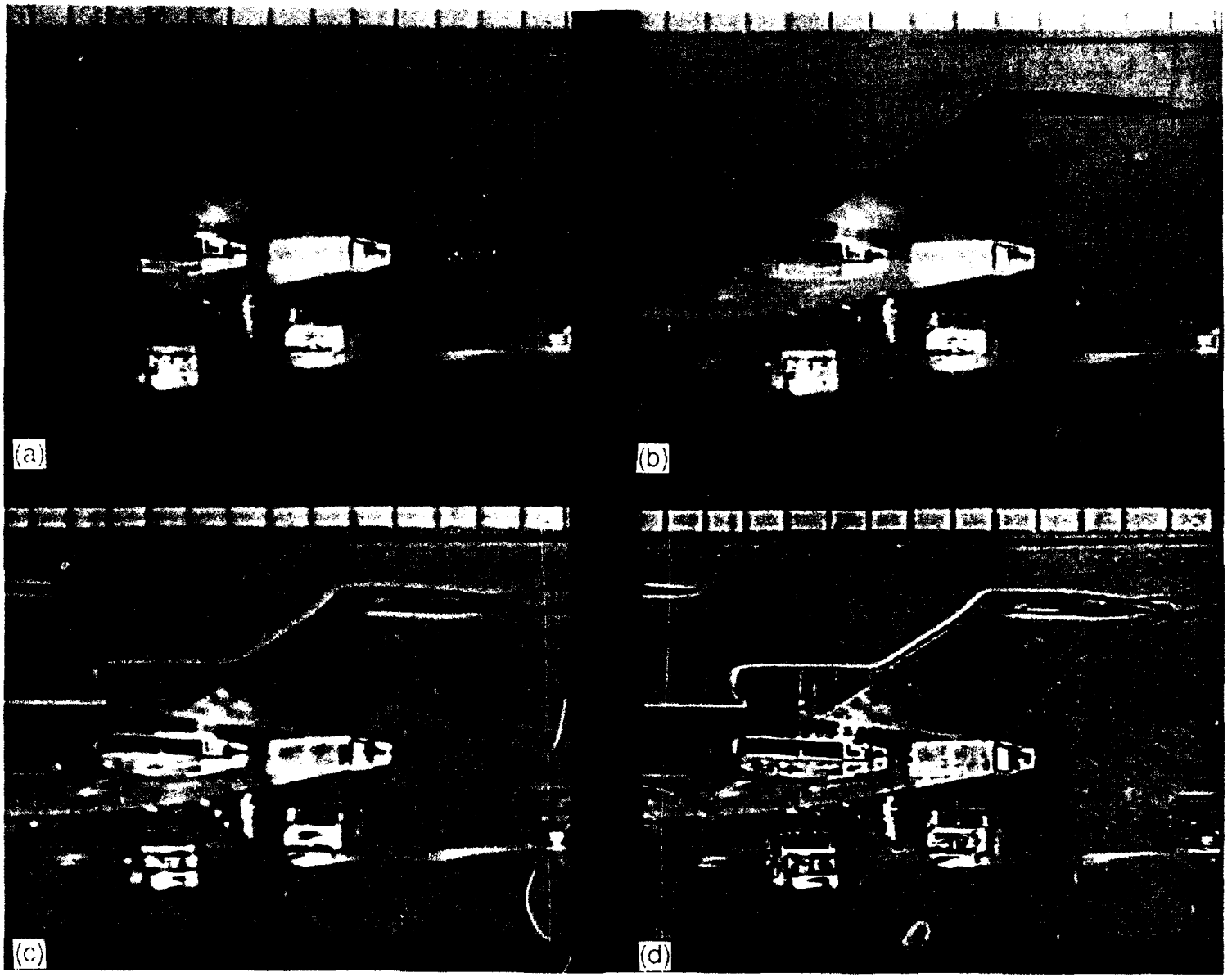



Figure 45. Examples of HP displays after high frequency enhancement with: (a) Eq. 8 mask; (b) Eq. 9, $a = 2$, $n = 3$; (c) Eq. 9, $a = 4$, $n = 9$ (WS); (d) MG mask.

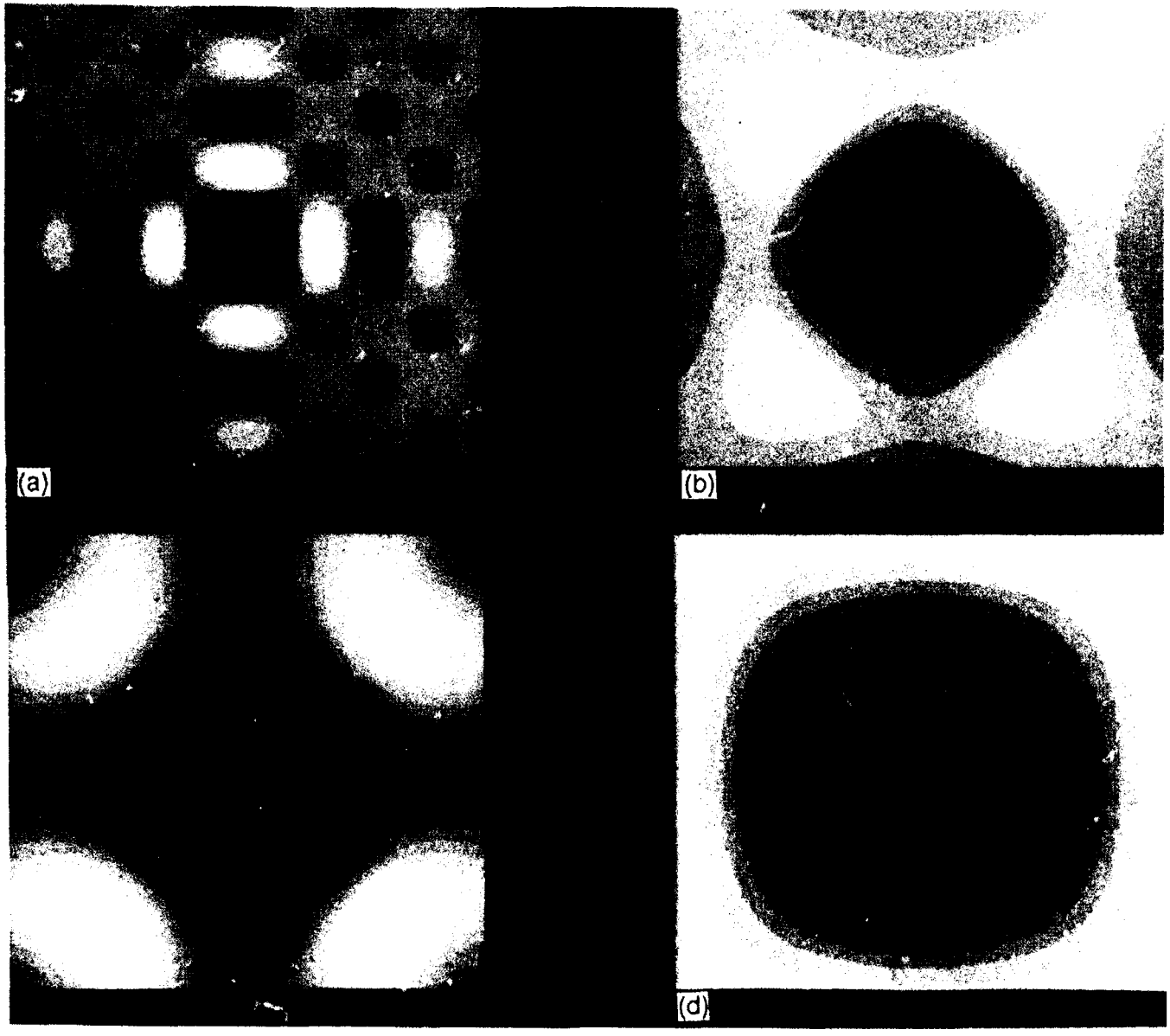



Figure 46. Filter transfer functions for: (a) WS; (b) SG; (c) MG; (d) WG.

$$G(r) = \begin{cases} 1 & \text{for } r = 0, \\ \lambda_1 - (\lambda_1 - \lambda_2)\exp(-r^2/\lambda_3^2) & \text{otherwise,} \end{cases} \tag{10}$$

where $r$ is the distance in frequency space from the (0, 0) frequency. Eq. 10 was strengthened and adjusted in the SFD by using it to filter the DFT of representative images and inverse-transforming the result to inspect the filtered image. The final Gaussian G was then transformed back into a (SD) convolution mask. We then approximated this mask using integers, shortened the mask extension, and toned down the effect slightly by trial and error. What finally emerged are the following three Gaussian convolution masks, referred to as strong, medium, and weak (SG, MG, WG) respectively:

$$SG: \quad \begin{bmatrix} 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & -3 & -2 & -1 \\ -2 & -3 & 37 & -3 & -2 \\ -1 & -2 & -3 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \end{bmatrix}, \tag{11}$$

$$MG: \quad \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 17 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}, \tag{12}$$

$$WG: \quad \begin{bmatrix} -1 & -2 & -1 \\ -2 & 13 & -2 \\ -1 & -2 & -1 \end{bmatrix}, \tag{13}$$

whose filter functions are compared to the starting WS filter in Figure 46b, c, d. The MG mask with power of two coefficients and the small 3 × 3 WG mask were designed for ease of implementation.

Figure 45d completes the sequence of sharpening comparisons of the airport scene with the use of the MG mask. Excellent local contrast enhancement is achieved with a more "natural look" than with the OP, RM or MP algorithms (see Figs. 33, 38, and 43). The DFTs of the airport displays in Figure 45b, c, d are shown in Figure 47. Comparison with Figure 44 indicates that the sharpening filters produce a more structured operation in frequency space, with greater de-emphasis of the low and mid-range spatial frequencies.

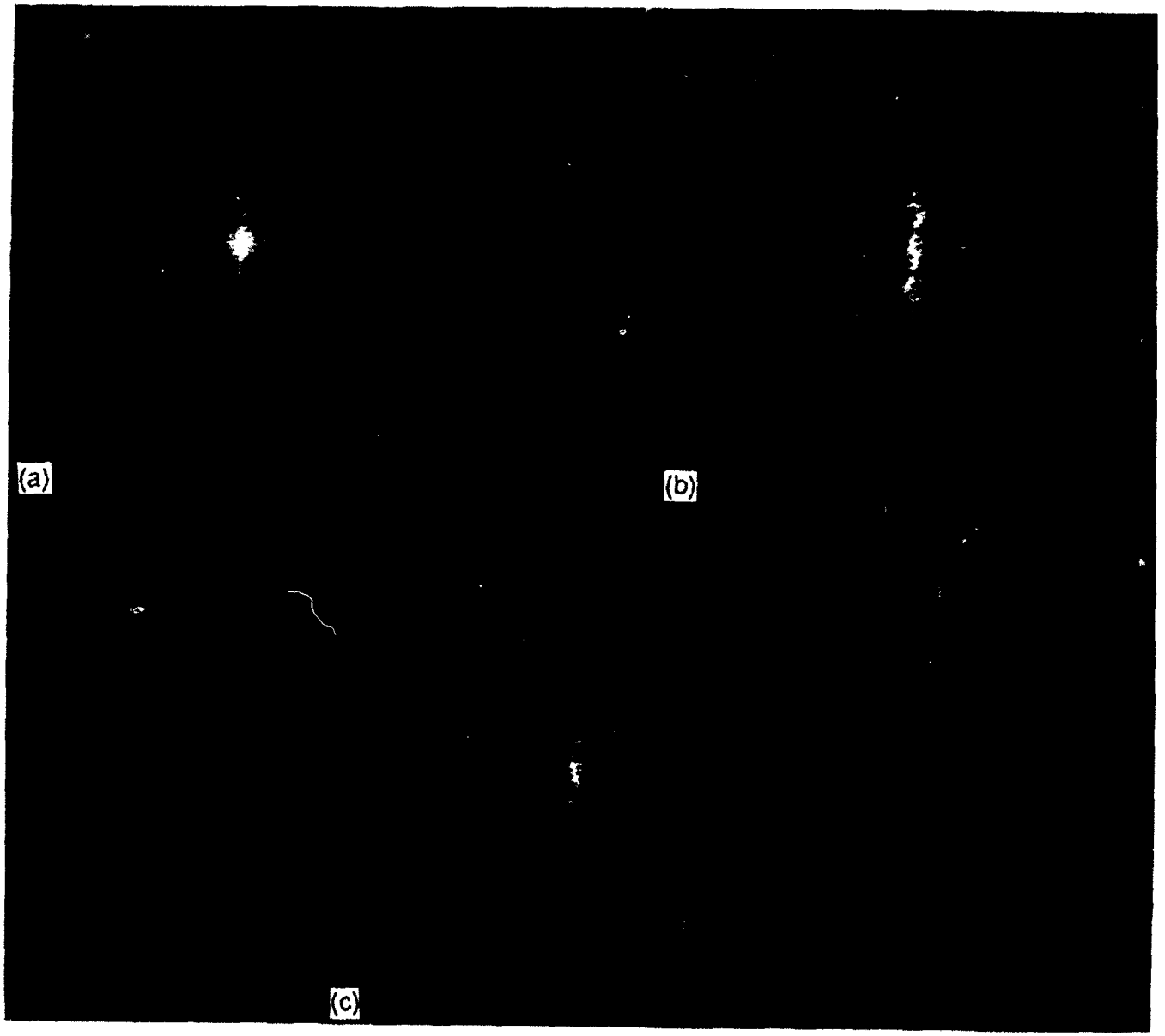



Figure 47. DFTs for airport displays after high frequency enhancement with: (a) Eq. 9, $a = 2$, $n = 3$; (b) WS; (c) MG.

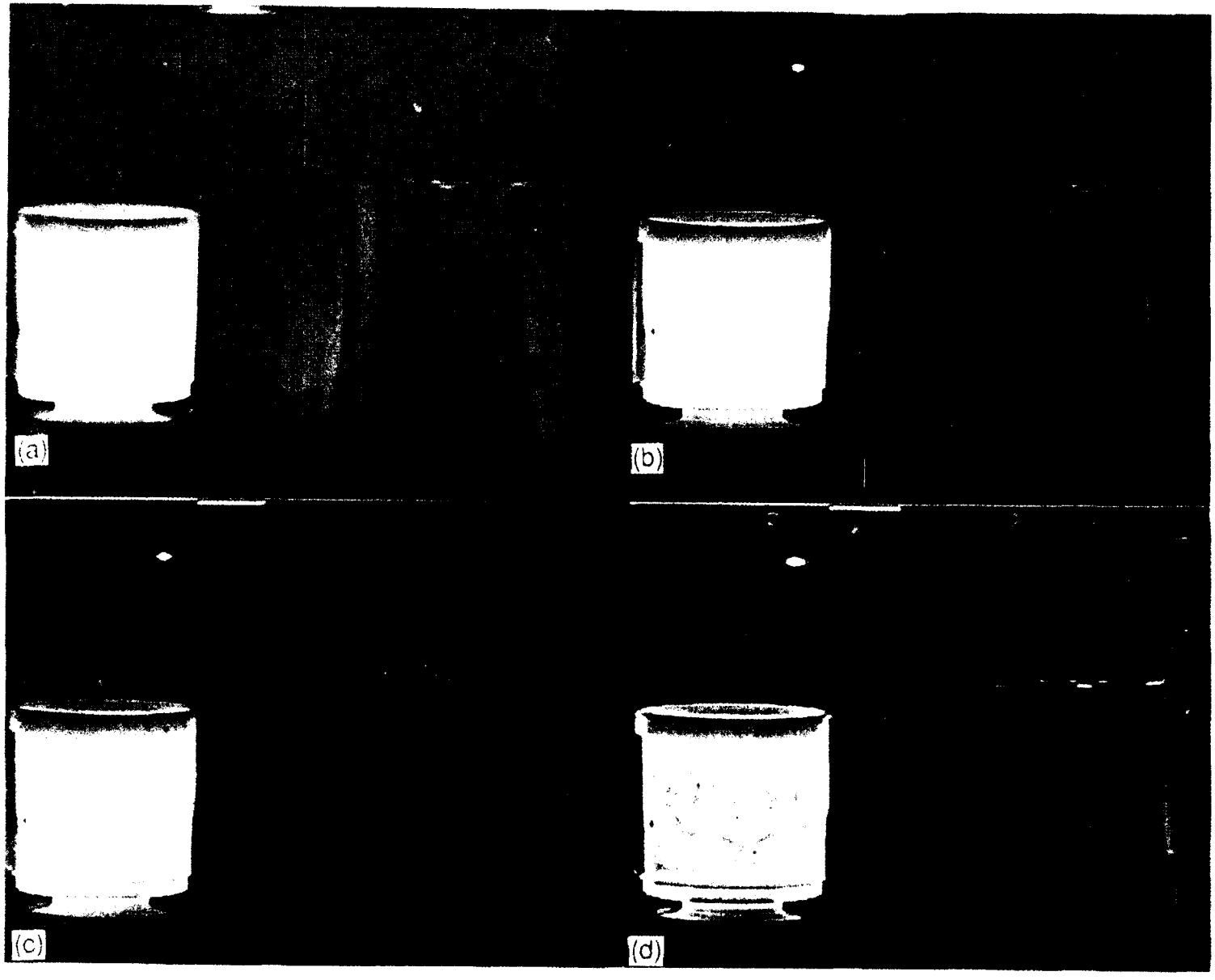



Figure 48. Examples of HP displays after high frequency enhancement with: (a) WS; (b) WG; (c) MG; (d) SG.

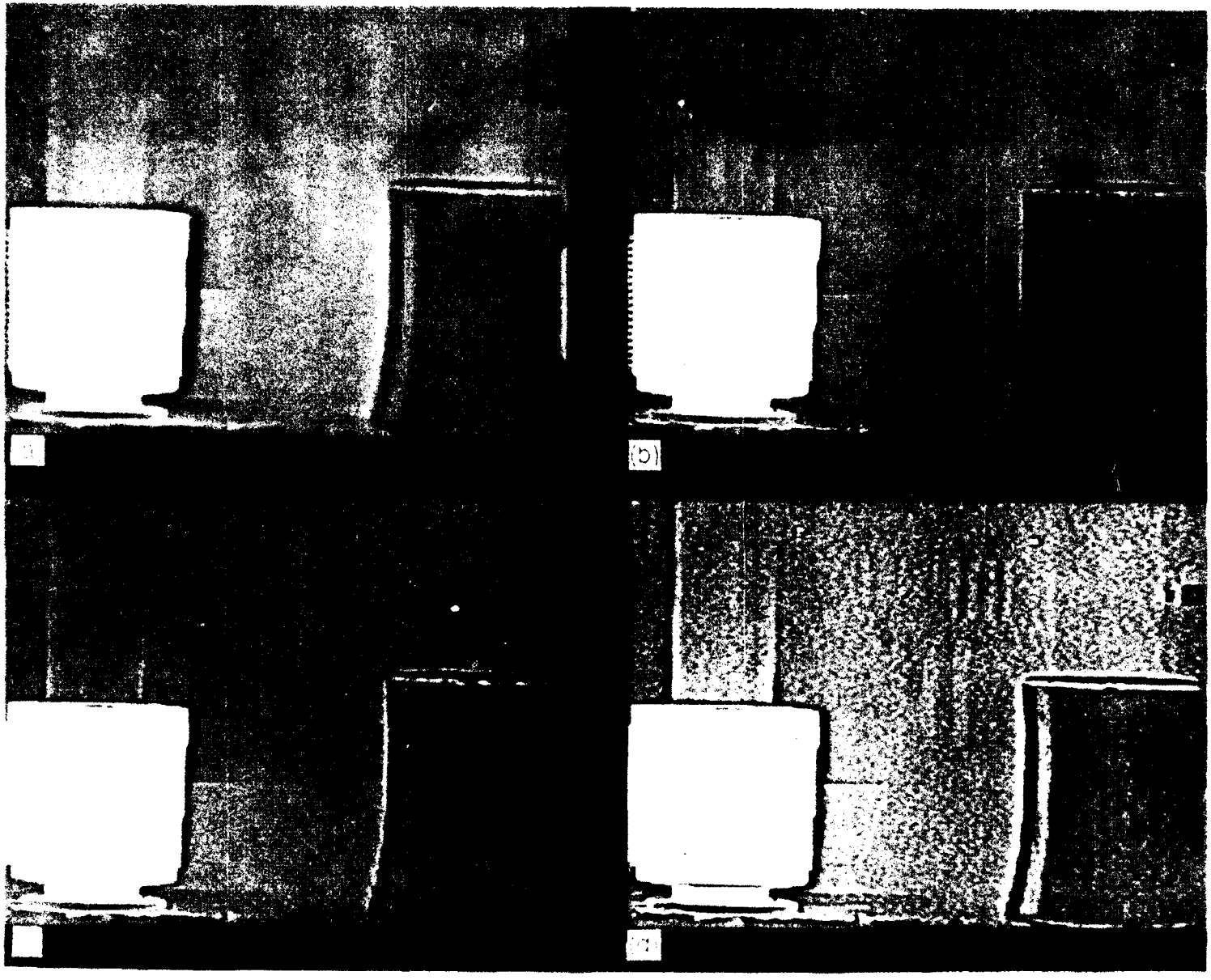



FIGURE 39. Examples of HP displays of noisy cups after high frequency enhancement with: (a) WS; (b) WG; (c) MG; (d) SG.

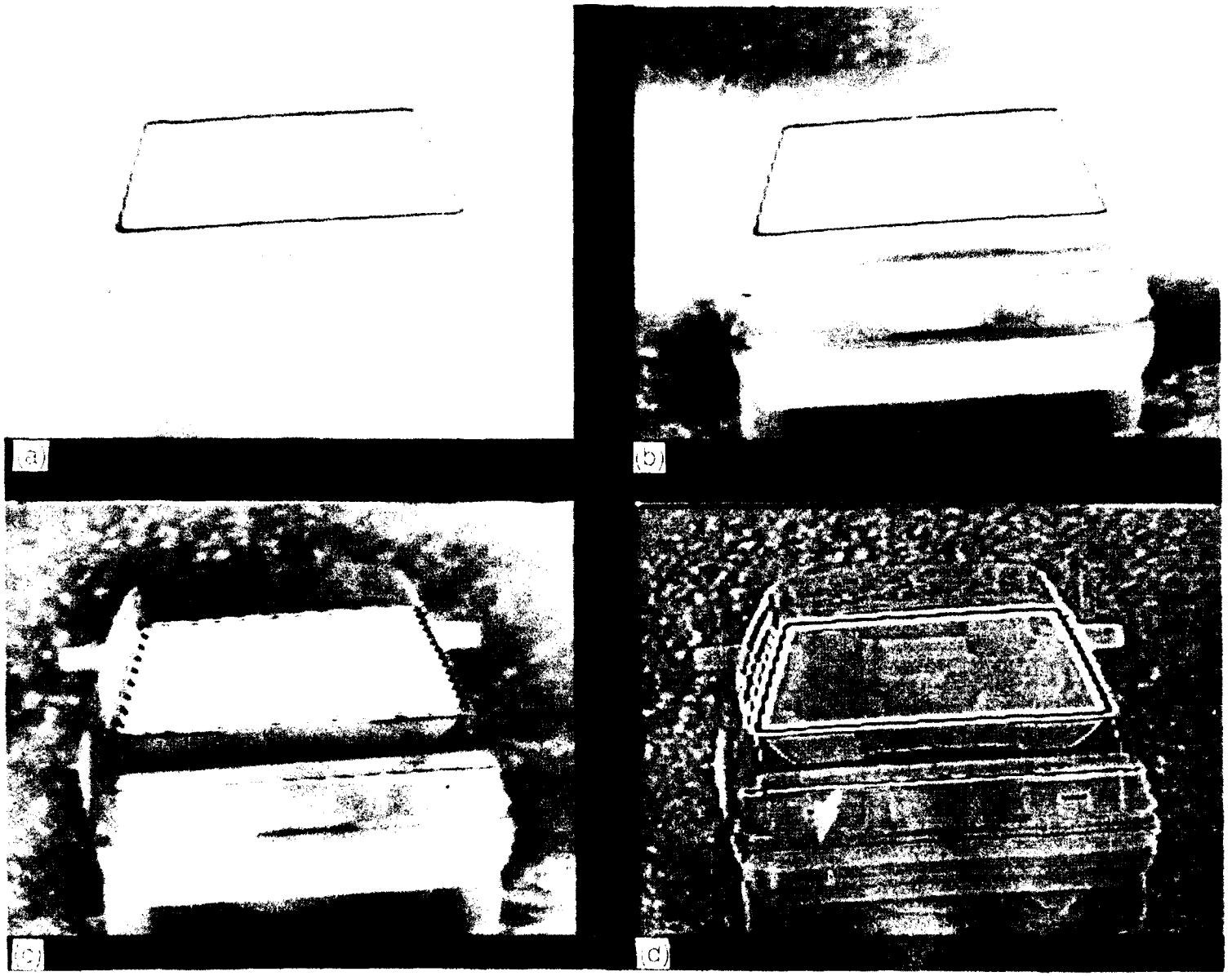



Figure 50. Comparison for bringing out license plate digits and car interior among: (a) HP; (b) OP; (c) RM; (d) SG/HP.

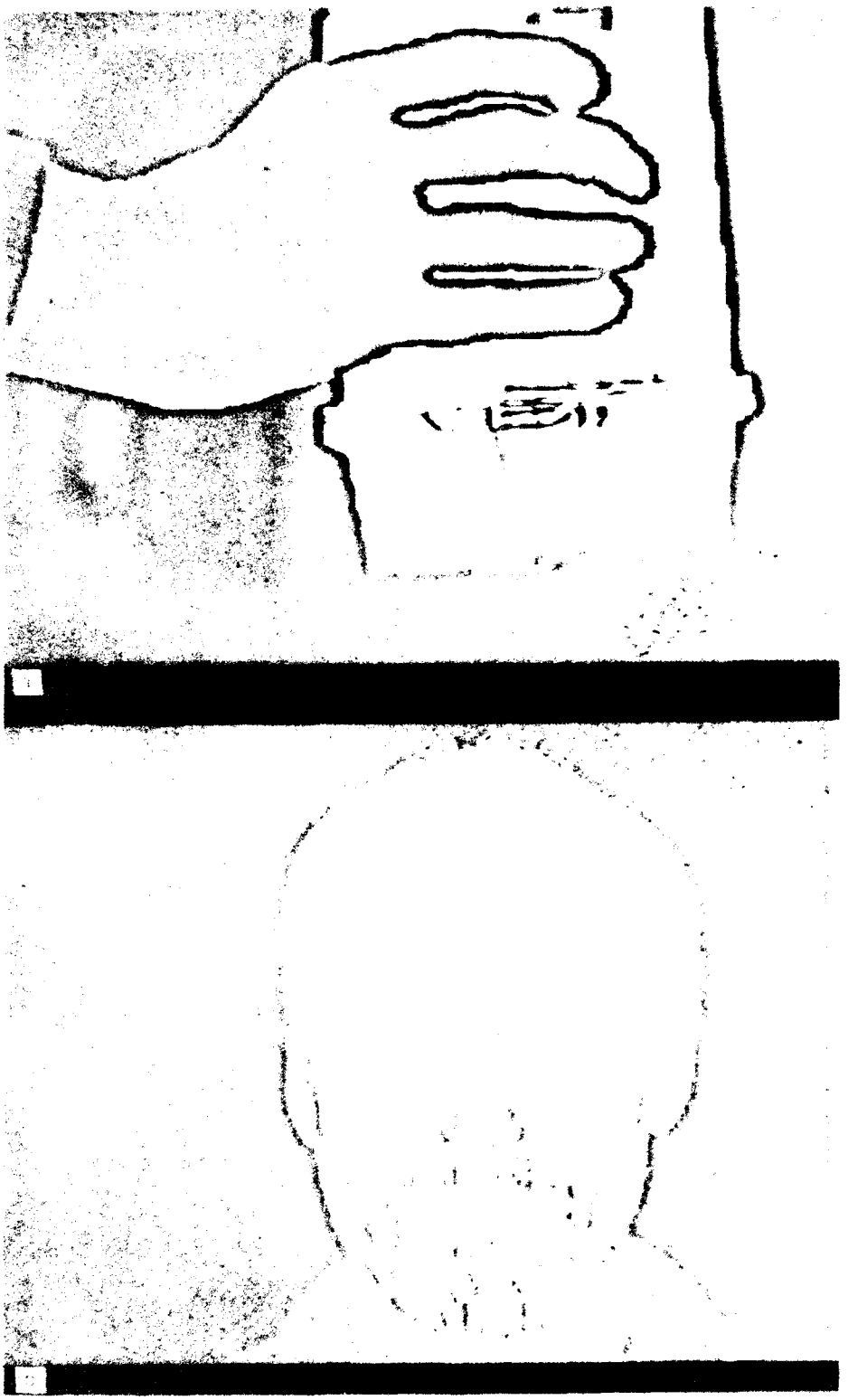


Figure 51. Examples of drawbacks of high frequency enhancement: (a) SGHP, (b) MGHP.

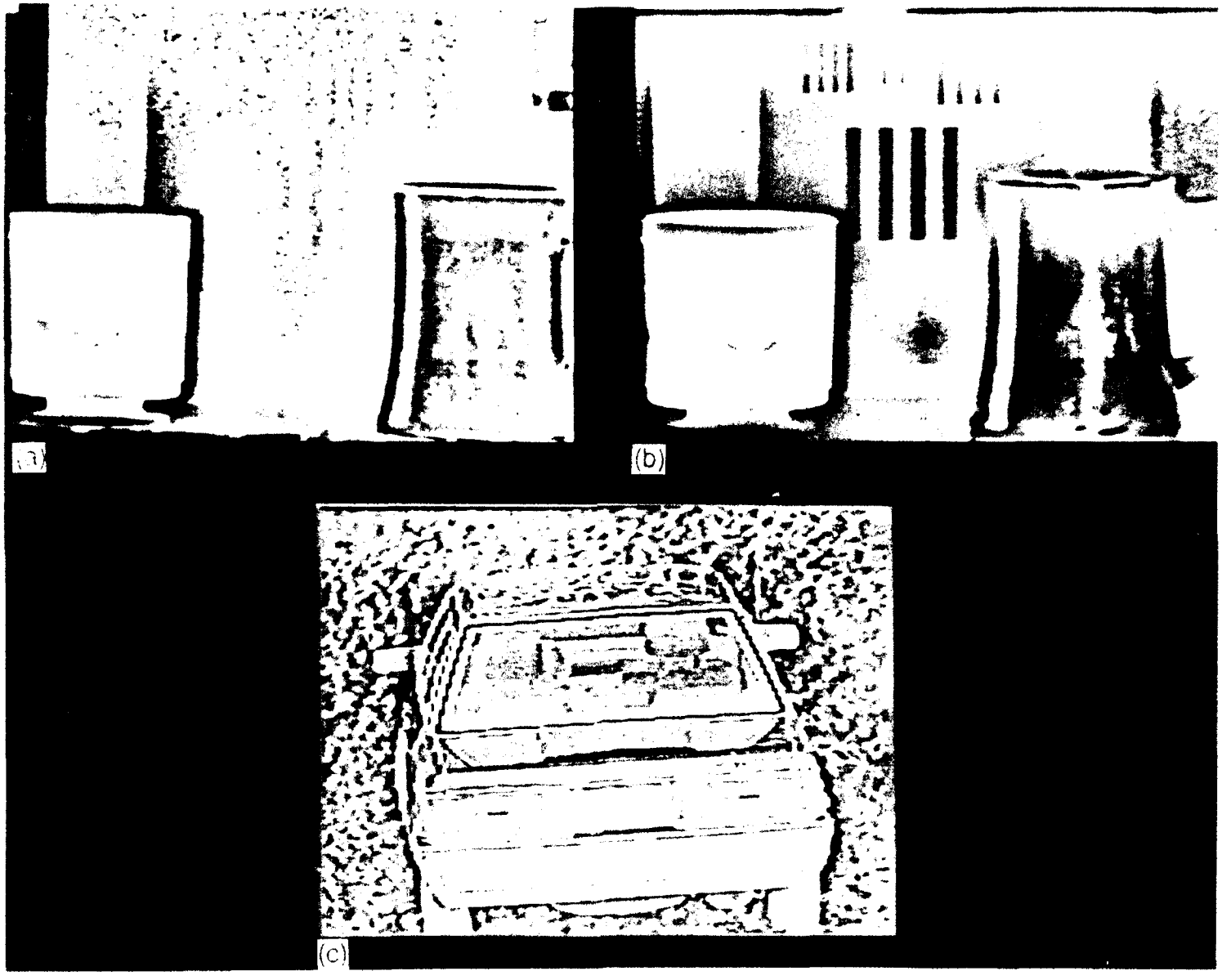



Figure 52. Examples of high frequency enhancement followed by OP: (a) WS/OP; (b) WS/OP; (c) SG/OP.

The WS filter ($a = 4$, $n = 9$ in Eq. 9) is compared in operation to the three Gaussian masks on both cup images in Figures 48 and 49. The weak or medium masks are often optimum, as in this case, for noisier images; while the SG mask, or even stronger versions, match up well with very low noise images. Especially for such low noise images, the Gaussian masks can reveal high frequency information missed by the previous algorithms, as for example in the license plate numbers on the automobile in Figure 50 (a 256-frame average and hence a low temporal noise image).

These masks – or for that matter any specific high frequency enhancement algorithm – have two drawbacks. First, they amplify both residual spatial noise (as in the horizontal lines on the warm cup in Fig. 49) and temporal noise (as in the bar patterns in the same image or the background in the face image of Fig. 51). Second, if the low contrast information is not high spatial frequency, such as the veins in the hand image (Fig. 51), modulo and local techniques (Figs. 42 and 34b) do a better job of enhancement.

A basic difference between the algorithms in this subsection and all previous algorithms should be underscored. We are here not mapping from raw signal to display values, but are rather preprocessing the raw signals. Hence we can join any of the previous algorithms in tandem with one of these high frequency enhancements. The examples shown so far have all been sharpening/HP. A very effective combination (results shown in Fig. 52) is to sharpen, e.g., with the WS algorithm, and then to map into an 8-bit display with the OP algorithm. This affords a stronger, more locally balanced, contrast enhancement than just OP, but with much reduced luminance artifacts (compare to Figs. 33 and 50). Apparently, the preprocessing with sharpening equilibrates to some degree the set of local numbers of occupied levels, giving smoother transitions between regions upon using the OP procedure (see 3.2).

## 3.5 Conclusions

The comments scattered throughout this section on the pros and cons of the various techniques for local contrast enhancement are gathered together and categorized in Table 2. We summarize some broad conclusions from this table.

The three types of local enhancement algorithms – local implementation of global algorithms, modulo processing, and high frequency sharpening – are all useful and sometimes complementary "software tools", which afford a comparable and effective degree of contrast enhancement, although one can find images or image types that match particularly well to each category. Local techniques like OP and SP are less predictable in their effects, with their circumstantial dependence on the number of locally occupied levels. The high frequency enhancement techniques rank high with respect to absence of artifacts, other than edge

**Table 2. Comparison of Algorithms for Local Contrast Enhancement**

| Algorithm | Strength of Enhancement | Natural Look/ Thermal Span | Main Artifacts | Ease of Real-time Implementation | Comments |
|---|---|---|---|---|---|
| OP | Volatile but usually strong | Fair | False luminance changes; volatile | Very difficult | "Volatile" due to image-dependent variation of local occupancy |
| SP | Very strong | Poor | Worse than OP, depending on window size | Extremely difficult | Windows 11 × 11, 21 × 21, 31 × 31, 41 × 41 |
| RM | Strong | Poor | Display sense reversals; gray scale arbitrary | Very simple | Raw signal compression and level shifts as useful adjuncts |
| MP | Moderate | Poor | Display sense reversals; gray scale arbitrary | Simple | Adapts to dynamic range of image |
| WS | Weak | Good | Edge overshoots | Moderate | |
| SG<br>MG<br>WG | Very strong<br>Strong<br>Moderate | Good | Edge overshoots | Moderate | Only enhances high freq. detail; amplifies temporal and residual spatial noise |

overshoots, and in "natural look". One might well ask, "What is a natural look for IR imagery?" We suspect that large regions which retain the monotonic thermal sense of the global mappings of Section 2 look more natural. Hence, modulo processing does not rank high in this respect and tends to look confusing, particularly to inexperienced observers.

If one turns to issues central to real-time operation in hardware on IR cameras, then RM processing, which entails simple global, signal-independent transformations, offers a computationally cheap form of contrast enhancement. Adding the two "bells and whistles" described in Section 3.3 (division of the raw signal by 2, 4, etc.; and additive shifts in the raw data scale) would increase the utility of such a camera adjunct.

## 4. HARDWARE IMPLEMENTATIONS AND FUTURE WORK

Many IR imaging applications such as low altitude night navigation, target tracking, and autonomous landing systems require a real-time automatic contrast control which provides optimized display imagery. While manual offset/gain adjustment on laboratory-designed cameras generally afford an excellent view of the IR scene content (in the hands of a skilled operator), frequent readjustment is required as the camera is panned or as the IR content of the scene changes. Real-time hardware implementations of the HP algorithm for this purpose have now been incorporated into several cameras designed in-house and afford a very useful alternative to the offset/gain controls.* In similar active areas of development, several US companies have already implemented HP or are now implementing it, typically in the UP or TP variation, for the same real-time function. So far, a fixed parameter in implementing UP or TP has been used. However, an implementation of UP with a programmable parameter (ranging say from HP itself to 1/32 under-sampling) would allow for some adjustments to the application by providing on-line flexibility in the dynamic range mapping.†

Aside from issues of computational complexity, other issues arise when real-time implementations are considered, such as interactions with temporal noise and mean display level flicker. A problem noticed in some of the HP implementations is the suspected presence of some frame-to-frame flicker noise, particularly in blander scenes (small total number of occupied levels $N$). Referring back to Eq. 5, we recall that the final display value of each pixel depends on its order number from 1 to $N$ in the hierarchy of occupied levels. Even in a

* The first such implementation was done in conjunction with the Hughes Aircraft Co., El Segundo, CA; details are available upon request from the authors.

† Such an implementation has now been done by the Eastman Kodak Co., Rochester, NY.

stable scene, noise variations from frame to frame, especially in their effect on sparsely occupied levels, can cause shifts in pixel order numbers. The most noticeable effects would arise from changes at the low or dark end (say as "detected" occupied levels disappear or are created) because such changes tend to affect the order number as a constant shift of a large majority of the pixels. The resulting rapid changes in the mean level of the display can become perceptible. Noise simulations have been performed and several fixes are being considered.[20]

We conclude this treatment of the display of IR images with a caveat. We have surveyed algorithms in two broad categories: "standard" algorithms used in other contexts (HE, LRM, homomorphic filtering, etc.) which we have tested on IR images; and algorithms newly devised for application to IR imagery (HP, OP, RM, etc.). The imagery driving our work and underlying this survey was based exclusively on PtSi staring technology and taken with cameras designed in our laboratory. More and more, IR imagery of this caliber is coming into the hands of foreign and domestic industrial companies which make PtSi cameras,[1] as well as imagery from other technologies such as InSb and HgCdTe. As staring IR imagery from other cameras, technologies, and wavelength regions (in particular, the 8-12 micron atmospheric window) becomes available, it would be desirable to revisit our surveyed algorithms. We therefore expect modifications and expansions of the perhaps somewhat parochial point of view of this report as more standard algorithms are tried on IR imagery or as new algorithms come along. Conversely, we anticipate possible use of some new algorithms such as HP, RM, or MP on other types of imagery with similar dynamic range requirements, such as medical imagery.

## ACKNOWLEDGEMENTS

We are grateful to Dr. J. M. Mooney, who contributed to much of the original work reported in this report, for his insightful reading of the manuscript. We are also grateful to Dr. F. D. Shepherd who has encouraged and supported our work since its inception. Finally, although the list is too long to mention individual names, we wish to acknowledge the many scientists past and present who worked in our laboratory on the conception, fabrication, design and development of PtSi IR cameras. Without their efforts, we would not have had this high quality imagery in our hands.

[20]Mooney, J. M., and Ewing, W. S. (1990). Private communication.

## REFERENCES


1. Shepherd, F. D. (1988). "Silicide Infrared Staring Sensors," Proceedings of SPIE, Orlando, Florida, 930, pp. 2-10.

2. Silverman, J., Mooney, J. M., and Vickers, V. E. (1990). Display of wide dynamic range infrared images from PtSi Schottky barrier cameras, Opt. Eng., 29: 97.

3. Silverman, J., Mooney, J. M., and Shepherd, F. D. (1991). Infrared video cameras, Sci. Amer., 266, No. 3, pp. 78-83.

4. Mooney, J. M., Shepherd, F. D., Ewing, W. S., Murguia, J. E., and Silverman, J. (1989). Responsivity nonuniformity limited performance of infrared staring cameras, Opt. Eng., 28: 1151.

5. Murguia, J. E., Mooney, J. M., and Ewing, W. S. (1990). Evaluation of a PtSi infrared camera, Opt. Eng., 29: 786.

6. Briggs, S. J. (1987). "Soft Copy Display of Electro-optical Imagery," Proceedings of SPIE, 762, pp. 153-170.

7. Briggs, S. J. (1981). Photometric technique for deriving a 'best gamma' for displays, Opt. Eng., 4: 651.

8. Schreiber, W. F. (1978). Image processing for quality improvement, Proc. IEEE, 66: 1640.

9. Woods, R. E., and Gonzalez, R. C. (1981). Real-time digital image enhancement, Proc. IEEE, 69: 643.

10. Pratt, W. K. (1978). Digital Image Processing, John Wiley & Sons, New York, pp. 307-344.

11. Hummel, R. (1977). Image enhancement by histogram transformation, Comp. Graph. & Imag. Proc., 6: 184.

12. Tom, V. T., and Wolfe, G. J. (1982). "Adaptive Histogram Equalization and its Applications," Proceedings of SPIE, 359, pp. 204-209.

13. Dion, D. F., and Cantella, M. J. (1984). Real-time dynamic range compression of electronic images, RCA Eng., 29: 42.

14. Silverman, J., and Mooney, J. M. (1988). "Processing of IR Images from PtSi Schottky Barrier Detector Arrays," Proceedings of SPIE, San Diego, California, 974, pp. 300-309.

15. Mooney, J. M. (1991). Effect of spatial noise on the minimum resolvable temperature of a staring sensor, Appl. Opt., 30: 3324.

16. Fahnestock, J. D., and Schowengerdt, R. A. (1983). Spatially variant contrast enhancement using local range modification, Opt. Eng., 22: 378.

17. Schowengerdt, R. A. (1983). Techniques for Image Processing and Classification in Remote Sensing, Academic Press, New York.

18. Dudgeon, D. E., and Mersereau, R. M. (1984). Multidimensional Digital Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey.

19. Wahl, F. M. (1987). Digital Image Signal Processing, Artech House, Norwood, Massachusetts.

20. Mooney, J. M., and Ewing, W. S. (1990). Private communication.

*U.S. GOVERNMENT PRINTING OFFICE:

***MISSION***

***OF***

***ROME LABORATORY***

*Rome Laboratory plans and executes an interdisciplinary program in research, development, test, and technology transition in support of Air Force Command, Control, Communications and Intelligence ($C^3I$) activities for all Air Force platforms. It also executes selected acquisition programs in several areas of expertise. Technical and engineering support within areas of competence is provided to ESD Program Offices (POs) and other ESD elements to perform effective acquisition of $C^3I$ systems. In addition, Rome Laboratory's technology supports other AFSC Product Divisions, the Air Force user community, and other DOD and non-DOD agencies. Rome Laboratory maintains technical competence and research programs in areas including, but not limited to, communications, command and control, battle management, intelligence information processing, computational sciences and software producibility, wide area surveillance/sensors, signal processing, solid state sciences, photonics, electromagnetic technology, superconductivity, and electronic reliability/maintainability and testability.*